ADM.
ARY
SEP 25 1934

VOLUME XXIX No. 187

JOURNAL
OF THE
AMERICAN STATISTICAL ASSOCIATION

SEPTEMBER, 1934

CONTENTS

SIGNIFICANCE OF INFANT MORTALITY RATES FOR SMALL GEOGRAPHIC AREAS. By Robert E. Chaddock 243

PITFALLS IN SAMPLING FOR POPULATION STUDY. By Clyde V. Kiser 250

RELIABILITY OF FOREST SURVEYS. By Bruce D. Mudgett and S. R. Gevorkiantz 257

NET EMIGRATION FROM THE UNITED STATES TO CANADA: 1909–1931. By Norman Leon Gold 282

RESEARCH AND STATISTICAL PROGRAM OF THE FEDERAL EMERGENCY RELIEF ADMINISTRATION. By Emerson Ross .. 288

FERTILITY OF FAMILIES ON RELIEF. By Samuel A. Stouffer 295

NOTES:

Boole's Challenge Problem. By Edwin B. Wilson 301

Index of Production of Manufactures Derived from Census Data, 1929, 1931. By Aryness Joy 305

On the Rapidity of the Decline in the Birth Rate. By P. K. Whelpton 308

London Meeting of the International Statistical Institute. By Walter F. Willcox 310

Taming the Weekly Seasonal. By G. M. Whitbright and M. Farrington 313

On a Certain Conclusion of Frisch's (316); Reply to Mr. Subramanian's Note (317); A Statistical Slide Rule (318); Corrections (318); Committee on Government Statistics and Information Services (319); Ought the United States to Resume Lending Abroad? (320); Technical Methods of Forecasting Stock Prices (323); Progress of the Census (325); Miscellaneous Notes (328); Members Added (337).

REVIEWS 339

PUBLISHED QUARTERLY BY THE
AMERICAN STATISTICAL ASSOCIATION
PUBLICATION OFFICE: Rumford Press, Concord, N. H.
EDITORIAL OFFICE: Columbia University, New York City

Price $1.50 per copy **$6.00 per annum**

NO. 187 (VOL. XXIX) SEPTEMBER, 1934

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

SIGNIFICANCE OF INFANT MORTALITY RATES FOR SMALL GEOGRAPHIC AREAS[1]

By Robert E. Chaddock, *Columbia University*

In a city of seven million people, with its great diversity of environmental conditions and population classes, it is important to have the vital statistics available by relatively small and homogeneous geographic units.

John Shaw Billings, in charge of the vital statistics of the 1890 Census, undertook for the first time to tabulate and publish the mortality of New York and a few other cities according to small geographic areas. This was not repeated until after the Census of 1910. The enumeration of population for New York City at that census was tabulated by "sanitary districts," now known as "census tracts." The Health Department soon began tabulation of the vital statistics of Manhattan Borough on the basis of "census tracts," to replace the old ward tabulations.

Most of the "census tracts," of which there are now over 3,000 in the Greater City, have been found to be too small for the computation of reliable rates. Therefore, they have been combined into what are known as "health areas." The Greater City has over 300 of these, averaging about 20,000 population each, which are now the official units of the Health Department for the tabulation of its vital statistics.

East Harlem Health Center was organized soon after the close of the Great War. It has proved to be a noteworthy experiment in coöperative health organization. For over a decade it was known as the East Harlem Health Demonstration, privately endowed and managed, but recently the City Health Department has assumed control of the Center. The original area included eight "census tracts," which have been combined into four "health areas," with a total of about 85,000

[1] Revision of a paper delivered at the Ninety-fifth Annual Meeting of the American Statistical Association in Philadelphia, December 29, 1933, at a session joint with the American Sociological Society.

population. The present Health Center District has been enlarged to include 20 "census tracts" combined into 10 "health areas," with almost a quarter of a million population.

We propose to examine the infant mortality rates of the health areas, especially in the original Demonstration District, in order to determine significant differences between them in respect to this sensitive index of health conditions. In order to increase the stability of the rates for each health area, the births and infant deaths for three years, 1929–31, have been included in the computations, Table I. The rates show wide

TABLE I

INFANT MORTALITY PER 1,000 BIRTHS, BY HEALTH AREAS, 1929–31, EAST HARLEM HEALTH CENTER DISTRICT, NEW YORK CITY

Health area	Infant rate 1929–31	Births for 3 years	Deaths for 3 years
17	71	1,516	108
21	59	1,134	67
22	58	1,799	104
26	68	1,331	91
17–26	64	5,780	370
20	99	2,846	283
25	79	2,105	166
28	66	1,269	84
29	70	1,358	95
30	94	1,403	132
33	86	1,435	123

variations in different parts of the entire district, from the low rates of areas 21 and 22 to the high incidence in areas 20 and 30. The Old Demonstration Area (17, 21, 22, 26) after a decade of experience in health organization, still had, in 1929–31, marked observed difference between the rates of areas 17 and 26 compared with those of areas 21 and 22. It is important to determine whether these observed differences between the rates of the areas in 1929–31 are real and persistent, and, if so, what are the causes? To what extent are the differences due to accidental, uncontrolled conditions?

When the rates, 1929–31, for the 81 health areas of Manhattan Borough are arranged in a frequency distribution the form is fairly symmetrical, Table II. The mean rate is 71.1 as compared with the median 70.3. The number of health areas in the East Harlem District (10) is too small to indicate the form of distribution of their rates.

There are recognized formulas for describing expected chance errors in rates and in their differences, but these formulas are based upon the assumption of the independent operation of the factors causing differences. Therefore, it is difficult to justify their use for describing errors

TABLE II

DISTRIBUTION OF INFANT MORTALITY RATES PER 1,000 BIRTHS, FOR 81 HEALTH AREAS OF MANHATTAN, NEW YORK CITY, 1929–31, AVERAGE RATES

Rate classes	Number of areas	
25 under 35	1	
35 " 45	5	
45 " 55	8	Mean equals 71.1
55 " 65	18	Median equals 70.3
65 " 75	16	Indicates slight skew
75 " 85	18	
85 " 95	6	
95 " 105	6	
105 " 115	1	
115 " 125	2	
	81	

in the rates of adjacent areas in which logically the factors causing differences would seem to be inter-related. If these formulas are employed to measure the chance errors in the differences of rates in the areas of East Harlem the results are as shown in Table III. There is

TABLE III

TESTS OF THE SIGNIFICANCE OF DIFFERENCES BETWEEN RATES OF HEALTH AREAS, 1929–31, EAST HARLEM HEALTH CENTER DISTRICT

N in Table = Not significant difference
S " " = Significant difference
? " " = Doubtful significance

Rate	Health areas									
99	20	S								
59	21	N	S							
58	22	N	S	N						
79	25	N	N	N	?					
68	26	N	S	N	N	N				
66	28	N	S	N	N	N	N			
70	29	N	S	N	N	N	N	N		
94	30	N	N	S	S	N	?	?	?	
86	33	N	N	?	S	N	N	N	N	N
Health areas		17	20	21	22	25	26	28	29	30
Rate		71	99	59	58	79	68	66	70	94

(1) Formula for expected variation in specific rate $\sigma_R = \sqrt{\frac{R(b-R)}{N}}$.

(2) Formula for expected variation in difference between two rates $\sigma_{(R_1-R_2)} = \sqrt{(\sigma_{R_1})^2 + (\sigma_{R_2})^2}$.

no assurance of significant differences in the rates for 1929–31, between any of the four areas of the Old Demonstration District (17, 21, 22, 26) or between any combination of them. Of course there is no doubt of significant differences between the rates of the Old Demonstration District and specific health areas of the larger present Health Center District (20, 30, and 33).

Since there are observed differences even between the rates of the four areas of the Old Demonstration District, and since there is doubt

as to the application of the usual error formulas to adjacent areas, we propose to resort to trend analysis for each health area in order to determine whether differences in the rates between the areas observed at one time persist over a longer period. Rates have been computed for each of the four areas (17, 21, 22, 26) and for each year since 1916. These have been plotted and a straight line fitted to the annual rates of each area, describing the trend of that area from 1921 to 1931 inclusive, Table IV. The rates for single years fluctuate more or less widely

TABLE IV

INFANT MORTALITY PER 1,000 BIRTHS, BY HEALTH AREAS AND FOR ENTIRE DEMONSTRATION AREA, 1916–1931, EAST HARLEM DEMONSTRATION AREA, NEW YORK CITY

(Annual rates 1916–31; Trends 1921–31)

Year	Area 17		Area 21		Area 22		Area 26		Entire demonstration area	
	Annual	Trend	Annual	Trend	Annual	Trend	Annual	Trend	Annual	Trend
1916..	109	....	117	....	97	....	106	....	106	....
1917..	111	....	92	....	105	....	85	....	98	....
1918..	112	....	103	....	92	....	105	....	102	....
1919..	90	....	95	....	72	....	86	....	84	....
1920..	126	....	118	....	99	....	109	....	112	....
1921..	79	87.6	59	73.9	78	76.2	90	83.2	77	81.5
1922..	87	86.4	90	72.4	76	74.0	89	82.3	85	79.7
1923..	89	85.2	76	70.9	66	71.8	65	81.4	74	77.9
1924..	90	84.0	67	69.4	71	69.6	89	80.5	80	76.1
1925..	94	82.8	58	67.9	57	67.4	62	79.6	68	74.3
1926..	66	81.6	67	66.4	74	65.2	90	78.7	75	72.5
1927..	61	80.4	63	64.9	65	63.0	61	77.8	63	70.7
1928..	114	79.2	73	63.4	58	60.8	90	76.9	84	68.9
1929..	81	78.0	68	61.9	65	58.6	93	76.0	76	67.1
1930..	67	76.8	52	60.4	63	56.4	62	75.1	62	65.3
1931..	70	75.6	57	58.9	44	54.2	50	*	54	63.5

Area 17—Slope = −1.2
" 21— " = −1.5
" 22— " = −2.2
" 26— " = − .9
Entire area— " = −1.8
* Trend fitted 1921–30.

about the trend values, as indicated by the diagrams (Charts I and II) and by Table V. These annual ups and downs suggest that any single

TABLE V

AVERAGE ANNUAL DEVIATION FROM TREND VALUES; BY HEALTH AREAS, 1921–31

Health Area 17	10.8 per 1,000
" " 21	7.1 " "
" " 22	5.3 " "
" " 26	12.8 " "

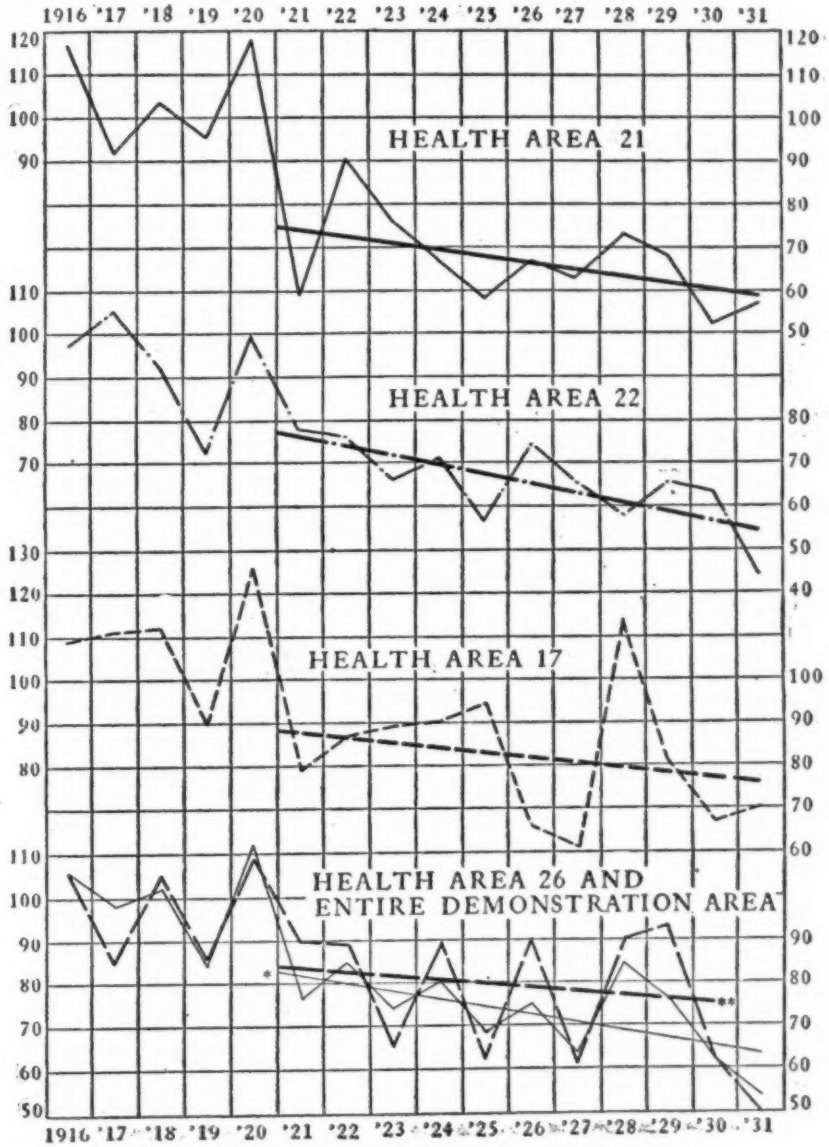
year, or even a combination of more than one year, are doubtful bases for comparative rates between adjacent areas, unless further analysis of persisting trends over as long a period as possible is made.

CHART I

INFANT MORTALITY PER 1,000 BIRTHS

FOR SPECIFIC HEALTH AREAS AND ENTIRE DEMONSTRATION AREA, 1916–1931
EAST HARLEM HEALTH DEMONSTRATION, NEW YORK CITY

ANNUAL RATES, 1916–31 STRAIGHT LINE TRENDS, 1921–31


1916 '17 '18 '19 '20 '21 '22 '23 '24 '25 '26 '27 '28 '29 '30 '31
HEALTH AREA 21
HEALTH AREA 22
HEALTH AREA 17
HEALTH AREA 26 AND
ENTIRE DEMONSTRATION AREA
*
**

* Light, continuous line represents the entire demonstration area.

** Trend fitted to 1921–30. Low rate of 1931 distorted slope of trend.

Chart II indicates that the trends of areas 21 and 22, which cross each other during the decade, are on about the same level. The trends of these two areas and of each of the other two areas persist throughout the periods on distinctly different levels. Other areas in the larger Health Center District may be treated in a similar manner, but the above analysis suggests that real and persisting differences between at least two of the groupings in the Old Demonstration District continue over the entire decade. What factors affecting the health of infants in these adjacent geographic units may account for the persisting differences?

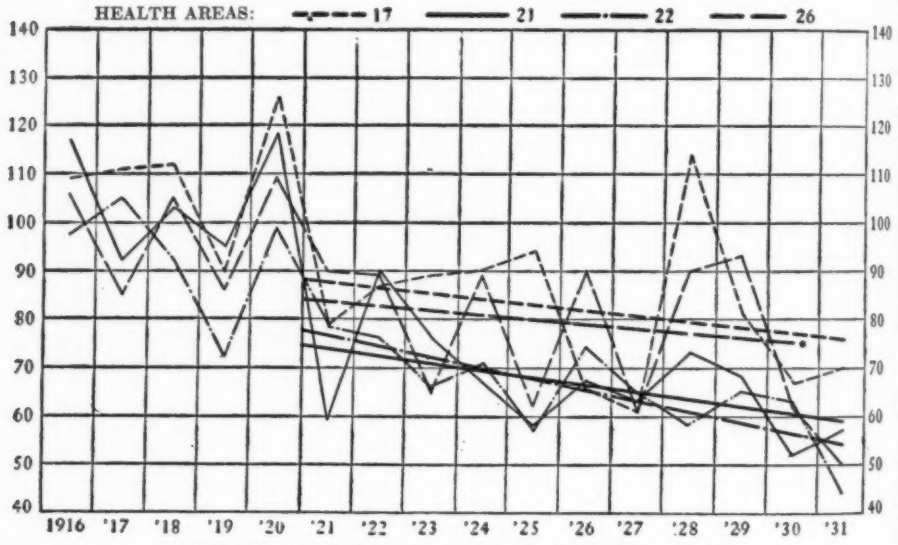
CHART II

INFANT MORTALITY PER 1,000 BIRTHS

FOR FOUR SPECIFIC HEALTH AREAS, 1916–1931

EAST HARLEM HEALTH DEMONSTRATION, NEW YORK CITY

ANNUAL RATES, 1916–31 STRAIGHT LINE TRENDS, 1921-31



* Trend fitted to 1921–30. Low rate of 1931 distorted slope of trend.

The initial approach to the analysis of vital statistics is made most directly through the use of geographic units. Health agencies are often organized for contact with people on a geographic basis. Birth and death records are filed at the Health Department by street address. The population upon which rates are based is tabulated by areas. But can we analyze the factors which cause apparent differences between the rates of adjacent areas by a further use of the data found in these small areas?

Having satisfied himself that real differences exist between the rates of adjacent areas, the investigator is interested in the problem of causation. If the statistical approach is employed in explanation it would seem necessary to abandon small geographic units for the moment. If the data from each of the small neighborhood areas are analyzed in order to set forth the possible factors affecting infant mortality—income levels, employment of mothers, nationality and race, method of feeding, pre-natal care, room-crowding, etc.—the cases of infant deaths in each category become so few that no conclusive results are possible from comparisons. Instability of rates based upon small numbers threatens the validity of explanations.

The alternative is to resort to analysis of mass data independent of particular geographic location, for the city as a whole or a large portion of it. Certain qualities of the population or of the environment characterize a neighborhood in such a manner as to affect the health of infants. The object of research is to determine what these factors are and how they are related to infant mortality. In this quest it is important to escape the hazard of drawing inferences from too few cases which is imposed by restricting the analysis to any narrow geographic unit.

With the knowledge of the important factors and their relation to the fluctuations in infant death rates, we return to the four health areas of the Old Demonstration District of East Harlem, and seek to account for the differences in rates observed to persist over the decade 1921–31. The investigator examines these areas from the viewpoint of differences in the specific characteristics of population or surroundings which, in the larger population of the city, have been shown to be associated with differences in infant mortality. This final analysis may serve to explain why the rate in one area differs from another, not in any one year but over a period of time.

PITFALLS IN SAMPLING FOR POPULATION STUDY

By Clyde V. Kiser, *Milbank Memorial Fund*

Adequacy and representativeness are fundamental requirements of reliable samples. Ross [1] has recently described the limitations of inadequate samples, especially when the variability of the data is great and many subdivisions required. The purpose of this paper is to describe certain difficulties encountered in securing a representative sample for the study of birth rates and to recount the steps taken to meet those difficulties.

The unsatisfactory nature of the original sample was due to the inception of the study in a morbidity survey. The original sickness survey was conducted in Syracuse in 1930 and 1931 by the United States Public Health Service in coöperation with the Milbank Memorial Fund and the Syracuse Department of Health. As first planned, data were also to be secured which could be used for a study of birth rates, but the prime considerations governing the collection of data were those concerning morbidity. In selecting areas to be studied, for instance, the aim was chiefly that of securing sections of widely different economic status in order that sickness comparisons might be made from that point of view.

Eight areas were selected for study and the original survey embraced 1,564 families from these areas. The smallest number of families enumerated for an area was 168 and the largest number was 256. The families enumerated comprised from 10 to 40 per cent of the total number residing in the census tracts approximating the districts selected.

In order to ascertain whether these schedules were suitable for population study or not, it was necessary to find out if the women involved were more or less fertile than the general run of women in the areas selected. This could be done most readily by comparing the sizes of households of the samples with those of the corresponding census tracts.[2] It was also possible to compare the samples with the census tracts in regard to certain factors associated with fertility, such as

[1] Frank A. Ross, (a) "On Generalisation from Limited Data," *Social Forces*, Vol. X, No. 1, October, 1931, pp. 32–37.

——— (b) "Ecology and the Statistical Method," *The American Journal of Sociology*, Vol. XXXVIII, No. 4, January, 1933, pp. 507–522.

[2] For the above comparison the census definition of household was adopted, ". . . a group of persons related either by blood or by marriage or adoption who live together as one household, usually sharing the same table."

The boundaries of census tracts did not coincide with those of the districts selected for study, but through combination of certain tracts it was possible to secure excellent approximation for several districts and at least fair approximation for any one.

nativity, country of birth of foreign born, and tenure, value, and rental of homes.

The above test revealed that the households included in the sample were larger than those of the areas in which they were located. This difference was not due to faulty selection so far as nativity, race stock, and economic status are concerned, as Tables I and II indicate. It

TABLE I

DISTRIBUTION OF HEADS OF HOUSEHOLD IN ORIGINAL AND FINAL SAMPLES AND IN CENSUS TRACTS ACCORDING TO NATIVITY AND OF FOREIGN BORN ACCORDING TO COUNTRY OF BIRTH

	Original sample		Final sample		Census tracts	
	Number	Per cent	Number	Per cent	Number	Per cent
Nativity						
Total white *	1,564	99.9	2,890	99.9	7,491	100.0
Native born						
Native parents	744	49.7	1,398	50.9	3,707	49.5
Foreign and mixed parents	415	27.6	761	27.7	2,092	27.9
Unknown parents	41	**	89	**	....	....
Foreign born	351	22.6	608	21.3	1,692	22.6
Unknown	13	....	34	....	....	....
Country of birth of foreign born						
Total	351	100.0	608	100.1	1,692	100.0
Canada	47	13.4	78	12.8	213	12.6
Ireland (Free State)	61	17.4	104	17.1	230	13.6
Eng., No. Ire., Wales, Scotland	43	12.3	77	12.7	244	14.4
Germany, Switzerland, Holland	74	21.1	139	22.9	261	15.4
Sweden, Norway, Denmark, Finland	5	1.4	8	1.3	30	1.8
Austria, Hungary	11	3.1	22	3.6	31	1.8
Italy	26	7.4	59	9.7	174	10.3
Russia and Balkan †	71	20.2	97	16.0	402	23.8
Other	13	3.7	24	4.0	107	6.3

* The fertility study was restricted to whites. In the total area studied there were, according to the 1930 census, only 95 Negro households and 9 of "other races," collectively only 1.4 per cent.
** Distributed proportionately in the above two percentages.
† Includes Poland, Czechoslovakia, Yugoslavia, Bulgaria, Roumania, Latvia, and Lithuania.

seemed probable that the selection of large families might be due to the failure of enumerators to revisit missed families, since married women who work away from home are more likely to be childless. The workers in the morbidity study had been primarily concerned with securing a sample representative of the area in regard to prevalence of sickness rather than size of household.

Another aspect of the original survey which was unsatisfactory for the study of birth rates was the fact that only one schedule was made for each household. Members of the household were included in the main roster but no provisions were made for securing fertility data for married individuals in the household other than the head and wife. This omission obviously was likely to constitute a selection so far as age is concerned, since the married women involved were generally daughters, daughters-in-law, mothers, and mothers-in-law of the head of the house.

The sins of our first sample, then, appeared to be those of omission. Plans were therefore made to secure the necessary fertility data from families called upon but missed altogether at the time of the sickness survey and from the married women other than wives of the head, i.e., secondary families, living in households already enumerated. In order to increase the sample somewhat further a few blocks or portions of blocks which had not been previously visited were enumerated. The further enumeration covered 1,069 additional primary families and 257

TABLE II

PERCENTAGE DISTRIBUTION OF HOMES IN SAMPLES AND CENSUS TRACTS ACCORDING TO TENURE, VALUE OF OWNED HOMES, AND RENTAL PER MONTH

Tenure, value owned homes, rental per month	Original sample	Final sample	Census tract
Tenure			
Total	100.0	100.0	100.0
Owners	52.8	50.5	53.3
Renters	47.2	49.5	46.7
Value			
Total	100.0	*	100.0
Under $3,000	3.5		1.8
$3,000– 4,999	13.9		5.5
5,000– 7,499	31.0		24.3
7,500– 9,999	25.2		26.4
10,000–14,999	18.4		26.0
15,000 and over	8.0		16.0
Rental			
Total	100.1	*	100.0
Under $20	4.6		4.6
$20–29.99	17.4		17.9
30–49.99	51.3		43.6
50–74.99	25.4		30.2
75 and over	1.4		3.7

* Value and rent data were not collected in the supplementary survey due to lapse of time since the depression and consequent incomparability with the original schedules and census data.

secondary families, an increase of about 85 per cent over the original number of schedules. Evening visiting and three or four repeated calls to the same home were often necessary but practically all of the families from whom data were desired were finally interviewed.[1]

The sample was generally improved by the supplementary survey. (See Table III, Chart I.) Further, in each district the distribution of households according to size was made remarkably similar to that based upon similar census tract data. This is readily seen in Chart II which shows deviations of the percentage distributions in the original and final samples from that of the census tract enumeration, plotted as zero. The differences that do exist are still in the direction of larger families in the samples than in the census tracts. The reason for this is

[1] One area (number 8) was not revisited for missed families because this was a slum section with a highly mobile and foreign population. Further, the data indicated that the sample drawn from this area was fairly representative.

TABLE III

DISTRIBUTION OF HOUSEHOLDS ACCORDING TO SIZE IN ORIGINAL AND FINAL SAMPLES, AND IN CENSUS TRACTS

Number in household	Original sample *		Final sample *		Census tracts *	
	Number	Per cent	Number	Per cent	Number	Per cent
Total	1,306	99.9	2,337	100.1	6,576	100.1
2	254	19.4	580	24.9	1,762	26.8
3	338	25.9	597	25.6	1,745	26.5
4	307	23.5	525	22.5	1,438	21.9
5	201	15.4	315	13.5	853	13.0
6	106	8.1	161	6.9	388	5.9
7	46	3.5	74	3.2	208	3.2
8	25	1.9	39	1.7	96	1.5
9 and over	29	2.2	42	1.8	86	1.3
Unknown	....	....	4	....	....	....

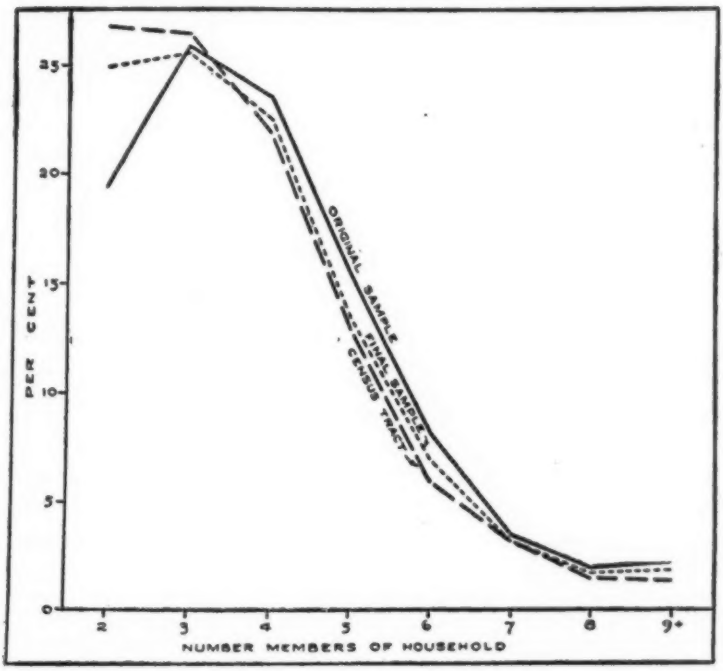
* District number eight and "one member" households were eliminated for the above comparisons.

probably the omission of households consisting exclusively of unmarried persons in the fertility survey.

Now that the representativeness of the sample is established, it is of more than academic interest to ascertain what differences exist between

CHART I

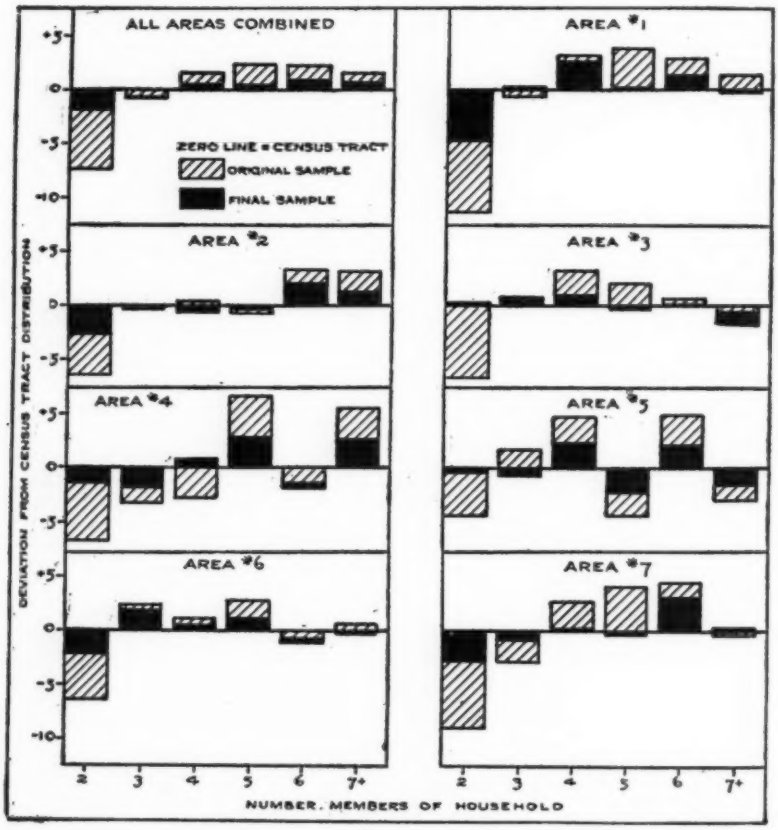
DISTRIBUTION OF HOUSEHOLDS ACCORDING TO SIZE IN ORIGINAL AND FINAL SAMPLES AND IN CORRESPONDING CENSUS TRACTS



women in the original primary families, additional primary families, and secondary families with respect to fertility and characteristics associated with fertility. Differences were found to be prominent with respect to fertility itself, age, and broken unions.

CHART II

HOW SUPPLEMENTARY CANVASSING IMPROVED THE ORIGINAL SAMPLE—DEVIATION OF SAMPLE DISTRIBUTION BY SIZE OF HOUSEHOLD FROM THOSE OF CORRESPONDING CENSUS TRACTS



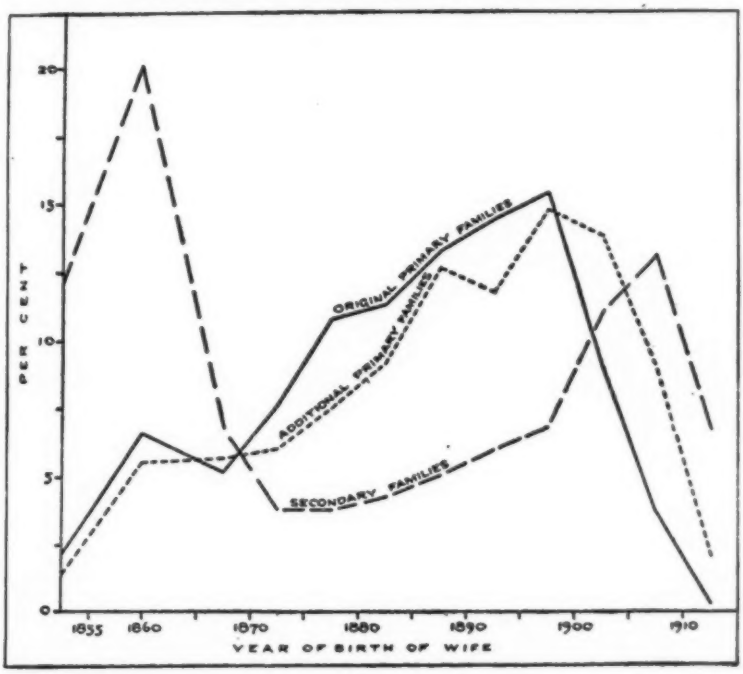
Actual birth rates (births per 100 wives) were computed separately for women in the three types of samples.[1] The crude rate was 239.5 for women in the original primary families enumerated, 210.9 for women in the additional primary families enumerated, and 194.5 for women

[1] The index "size of household" was not only too indefinite for this internal comparison, but in the nature of the census definition of household, could not be used for the comparison of primary and secondary families with respect to fertility.

in the secondary families.[1] While there are interesting differences in age and incidence of broken marriages among women in the three types of samples, the differences persist when these factors are held constant. When standardized for age [2] respective rates of women in the above types of samples are 234.2, 213.0, and 181.3. Further, our data indicate that the above differences in fertility cannot be attributed to differences in social class composition.

CHART III

DISTRIBUTION OF WIVES ACCORDING TO YEAR OF BIRTH IN ORIGINAL PRIMARY, ADDITIONAL PRIMARY, AND SECONDARY SAMPLES



Differences in the three types of samples with respect to the proportions of women who are childless also illustrate the necessity of recalls and the inclusion of secondary families in the sample. Considering women of all ages, 18.6 per cent of those in the original primary sample, 23.4 of those in the additional primary sample, and 28.5 of those in the

[1] District number 8 was eliminated altogether for this comparison because no attempt was made to enumerate families originally missed in that area.

[2] The standard used here is the age distribution of married, widowed, and divorced white female residents of Syracuse, as reported by the 1930 Census.

In coding age of wives not living with their husbands, the age at separation was taken if the separation occurred before the wife reached 45 years of age.

secondary sample were childless. The greatest difference between the samples was found among women 30–44. In this age group, only 15.3 per cent of women in the original primary sample were childless as compared with 22.3 per cent of those in the additional primary, and 34.4 per cent of those in the secondary sample. The small differential that does exist among women 45 years of age and over is in the opposite direction. This reversal is to be expected because the older women of the secondary families are largely mothers or mothers-in-law of the head of the household.

The selective factor involved in failure to make recalls and in failure to enumerate secondary families is also manifested in two attributes associated with fertility—age and broken unions. The women in the additional primary samples were a little younger than those in the original sample, but the age curves follow the same pattern. (See Chart III.) The age curve for women in the secondary sample is strikingly different, a peak being shown at either extreme of age. This is due to the fact that the secondary sample is composed largely of children or parents of the heads of the main households.

The women included in the secondary families were to a greater degree than those in primary families, survivors of broken unions. Many were doubtless living in the households of their parents or children simply because their own homes were broken. In only 47.2 per cent of the cases were the women of the secondary families living with their husbands at the time of the enumeration. The corresponding ratio for women in the original primary sample (exclusive of district number 8) was 89.1 per cent and for women in the additional primary families 84.8 per cent. Similar differences were found to be consistent in each of the districts considered separately.

The broad implications of this report are that careful study must be made of factors likely to influence the representativeness of a sample and that the safeguards adopted must depend upon the nature of the problem. So far as samples for population study are concerned, a definite bias is introduced unless the sample is selected with care and, once selected, is thoroughly canvassed. It is a persistent bias which cannot be corrected by mere extension of the sample. Revisits to homes originally missed and the inclusion of secondary families in the sample are basic to representative sampling for studies of birth rates.

RELIABILITY OF FOREST SURVEYS

By Bruce D. Mudgett[1] and S. R. Gevorkiantz[2]

The Federal Government has recently made provision, by the passage of the Sweeney-McNary Forest Research Act, for a survey of the nation's forest resources; and preliminary surveys have already been made for the purpose of laying down methods of procedure to be followed in the larger task. This paper deals mainly with the interpretation of the results in one of these preliminary surveys, that of Washburn County, Wisconsin, made during the summer of 1931 by the University of Wisconsin, the Wisconsin Conservation Department and the Lake States Forest Experiment Station.

This survey covered twenty-four townships, an area of about 553,687 acres. It is rectangular in shape, twenty-four miles wide and thirty-six miles long. The county lies in a transition belt between the northern hardwood forest and the central oak-hickory forests. The soils of the county are prevailingly sandy although certain heavy soils occur in the southern part and in the extreme northwest. Owing to its glaciated topography and especially to recent cutting, burning and other disturbing factors, the forest types in the county are rather small in extent and are not well distributed throughout the area.

Of the various data collected in this and other surveys similar to it, we shall be particularly interested in two items: (1) the area of each type of forest cover and (2) the volume of timber by species, in the total territory covered by the survey. Methods used in the survey will be studied for the purpose of determining the degree of reliability that can be placed upon estimates of area and volume.

Methods of obtaining survey data.—Ideally, here as elsewhere, a complete record of the facts is desired—an exact measurement of area of each forest type and a count of all trees by sizes; but actually the ideal cannot be realized. A complete forest census would carry the expense beyond all practical limits and it is necessary to use methods of estimation based upon a partial survey. There follows the necessity of judging the degree of error of estimated magnitudes, the need to determine within what range, on either side of an estimated area or volume, the unknown true value probably lies.

Both in the method of taking the partial census and in judging the reliability of the final result the problem is a statistical one. The

[1] University of Minnesota.
[2] Lake States Forest Experiment Station, Forest Service, United States Department of Agriculture.

theory of sampling now comprises a considerable body of statistical method that may be appealed to in laying out the method of partial surveys and in interpreting their results. The theory in its pure aspects deals with hypothetical or idealized situations (with the perfect dice and perfect coins of our imaginations) and its application to real situations is always more or less imperfect.

The data upon which forest areas and volume of timber have been estimated have been obtained in preliminary surveys by the following procedure. Parallel survey lines have been run across the area to be surveyed at intervals of one, two, or three miles and on each line small rectangular plots, of one-quarter acre or other size, have been laid out at short intervals (e.g. eight per mile) along the survey lines. On these plots a complete census has been taken of all information desired. The estimated proportion of area occupied by any given forest type has then been obtained by taking the ratio of the number of plots covered by this forest type to the total number of plots. This is known as the plot-count method of estimation. A second method is used in making area estimates by recording the distance traveled through each forest type along the survey lines and taking the ratio of this distance to total length of lines run as the estimated proportion of the given forest cover. This is called the linear measurement method of estimation.

The theory of Bernoullian, Poisson and Lexian distributions.—The question at issue is, How accurate is such an estimate of area? The pure theory of sampling for attributes provides estimates of reliability of a proportion obtained under various hypothetical conditions, and it becomes a problem then of judging what set of conditions in pure theory are most nearly approximated by the sample as actually taken. The pronouncements of pure theory may even go beyond judgment of this sample, for it may suggest better methods of taking the actual sample than that used.

The part of the pure theory of sampling here needed is developed in a number of standard texts on statistics [1] to which the unfamiliar reader is referred. Only the simplest description is given below, together with the necessary formulas.

Let there be a set of urns, U_t, $(t, 1, 2, \ldots . s)$, each containing balls of identical kind except that some are black and some white; the number of balls in urn (t) is n_t and the proportion of white balls, p_t. Thus for the several urns:

Urns:	U_1, U_2,	. . . .	U_t,	. . . .	U_s
Proportion of white balls:	p_1, p_2,	. . . .	p_t,	. . . .	p_s

[1] See Yule: *Introduction to the Theory of Statistics;* Rietz (ed.): *Mathematical Handbook of Statistics;* Arne Fisher: *Mathematical Theory of Probabilities.*

Number of balls: $N=n_1 + n_2 + \ldots + n_t + \ldots + n_s$

whence $Np_0=n_1p_1+n_2p_2+ \ldots + n_tp_t+ \ldots + n_sp_s$

p_0 being the proportion of white balls in all urns combined.

The problem is to estimate the proportion of white balls, p_0, in the whole set of urns by a partial drawing or sampling instead of a complete census count of all the balls. In drawing any ball from any urn one condition shall be satisfied—that every ball in the urn shall have an equal chance of being selected. This condition will be satisfied if, in drawing n balls from urn t, the contents of the urn are thoroughly mixed, then one ball is drawn and its color noted; and if, on the second and succeeding draws, the ball previously drawn is replaced in the urn and the whole contents thoroughly mixed again before the succeeding draw takes place.

A sample of n balls is to be drawn and the proportion of white balls in this sample is to be used as an estimate of the proportion of white in all the urns combined (the whole universe). Three methods of selecting the sample are available. The contents of the several urns may be combined and thoroughly mixed and then a sample of n drawn at random. The results of repeated drawings of this sort will produce a Bernoullian frequency distribution with a known standard deviation and the latter can be used as the error of estimate of such a sample. On the other hand, a sample of n may be drawn by selecting proportionately from the several urns, the drawings from each urn being random, kn_1 from U_1, kn_2 from U_2, ... and kn_s from U_s; where $kn_1+kn_2+ \ldots + kn_s=kN=n$. The results of repeated drawings of this sort will produce a Poisson frequency distribution and its standard deviation, which is known, can be used as the error of estimate of a proportion obtained by taking a stratified, proportional, or Poisson sample. The third method is to select an entire sample of n from a given urn, the drawing to be random, and to use the proportion observed in this sample as an estimate of the true proportion in the universe. If r such samples of n are taken so that kn_1 samples come from U_1, kn_2 from U_2, ... and kn_s from U_s, where $kn_1+kn_2+ \ldots + kn_s=r$, the results will give a Lexian frequency distribution, and its standard deviation, which is known, can be used as the error of estimate of a Lexian sample.

In each of these three methods of taking samples, Bernoullian, Poisson and Lexian, the results of a given sample will in general not give a proportion in agreement with p_0 but will give $p=p_0+e$; a second sample will give $p'=p_0+e'$, and so on. The three frequency distributions briefly described above give, each for its appropriate method of

taking the sample, the error curve of the above e; so that for each of these three methods of taking samples the standard error of estimate is known. These will be designated σ_B, σ_P and σ_L. The three formulas are given herewith:

$$(1)\quad \sigma_B^2=\frac{pq}{n}$$

$$(2)\quad \sigma_P^2=\sigma_B^2-\frac{1}{n}\cdot\sigma_p^2$$

$$(3)\quad \sigma_L^2=\sigma_B^2+\frac{n-1}{n}\cdot\sigma_p^2$$

$$=\sigma_B^2+\sigma_p^2 \text{ if } n \text{ is large.}$$

Corresponding formulas for *numbers* of successes may be obtained by multiplying the above results by n^2, but the formulas for proportions are more frequently used in practice. The σ_p which appears in formulas (2) and (3) is the standard deviation of the proportions, p_i, in the several strata (urns), obtained from the formula:

$$\sigma_p^2=\frac{\Sigma n_i p_i^2}{N}-p_0^2.$$

The results show that the error of estimate of a Poisson sample is least among the three and that that of a Lexian sample is greatest. If the universe being investigated is not homogeneous, that is, if stratification is present so that the frequency of the attribute under investigation varies from stratum to stratum, then the frequent admonition that a sample must be a random sample is mistaken, for the most accurate estimate possible is desired from the sample; and if stratification is present and its character known, a stratified or proportional sample will have a smaller error of estimate than a random sample. On the other hand, if stratification were suspected but its character not fully known, it is possible that the sampling procedure might succeed only in representing a portion of the strata. In this case the sampling procedure would have satisfied Lexian conditions and the error of estimate would then be greater than for random sampling. Note also that the effect of Lexian conditions is much more potent than that of Poisson conditions, for the Lexian variance is increased over Bernoullian by the variance of the proportions, p_i, while the Poisson variance is decreased as compared with Bernoullian by only one-nth part of the variance of p.

The practical problem of sampling where stratification is present or is suspected, is to take a Poisson sample if possible and to estimate its

error by the Poisson formula. If this is impossible and if Lexian conditions of sampling have been introduced willy-nilly the error of estimate of the sample should be based on the Lexian formula. These considerations are of fundamental importance in estimating area and volume of timber in forest surveys.

ESTIMATES OF AREA AND THEIR RELIABILITY

This brief statement of the theory of Bernoullian, Poisson and Lexian dispersion is the basis in statistical theory for answering the question, how accurate is an estimate of forest area made by the plot count or by the linear measurement method. The first step in the answer is to determine whether the *universe* which was surveyed classifies, to a close approximation, as a homogeneous universe or whether it is stratified. To use the analogy of the urn and balls, is the forest area that is being surveyed similar in character to a single urn with given proportions of black and white balls, that is, with a fairly even distribution of a given forest type over the entire area; or is it better described by a series of urns with differing proportions of white balls, that is, more of the given forest type in one part of the area, less of it in another?

Stratification revealed by survey lines.—One might almost assume with safety that it is of the latter type, that the forest area is stratified with respect to any particular type. But we shall see what evidence is given by actual surveys. The Washburn County, Wisconsin, survey ran eighteen survey lines across the county spaced at intervals of two miles. Sample plots two chains long and one-half chain wide (=one-tenth acre) were laid out at ten-chain intervals along these lines, giving eight plots to the mile. Area estimates in the survey were made by the plot-count method. The actual number of plots of northern hardwoods found on these eighteen lines were, taking the lines consecutively, 26, 45, 37, 17, 40, 7, 10, 0, 5, 5, 0, 0, 5, 2, 13, 26, 25, 11, a total of 274. The question is whether the variation from line to line is such as can be accounted errors of random sampling in a homogeneous universe or whether it varies to a significant degree from the latter. To test this, the best estimate of p_0, the true proportion in the entire area, which is available is the calculated value of p for the entire eighteen lines, namely, .07928. Each line containing 192 plots, the standard error of sampling of a random selection of 192 is $\sigma_B = \sqrt{npq} = 3.7437$.

The calculation of the standard deviation of the eighteen values given above gives $\sigma = 14.156$, a much larger value than σ_B. A test of significance is available by use of the formula for the standard error of sampling of a standard deviation in a binomial, or Bernoullian universe, viz.,

$$\sigma_\sigma^2=\frac{1}{4N}\left(2npq+1-6pq\right)=.397110967$$

$$\sigma_\sigma=.63016$$

The N in the formula equals 18, the number of lines. If the actual standard deviation does not differ from σ_B by more than two or three times the value of the standard error of σ_B it can be said to have arisen as a fluctuation of random sampling. But

$$\frac{\sigma-\sigma_B}{\sigma_\sigma}=\frac{14.156-3.744}{.6302}=16.5$$

a difference that makes the dispersion of these eighteen values non-Bernoullian as a practical certainty. It is also interesting to compare the standard deviation of these values with a Lexian error of sampling:

$$\sigma_L^2=npq+n(n-1)\sigma_p^2=213.36632$$

$$\sigma_L=14.61$$

$$\sigma=14.16$$

a result readily forecast when it is recalled that the Lexian sample in this case is the record of any given line and that the standard deviation of a distribution of such samples is $n\sigma_p$ calculated from data of the eighteen lines.

The above analysis is used only to show that the forest area is stratified. The same results appear with respect to all of the different cover types in the Washburn County survey—the different survey lines represent Lexian conditions, or different urns with varying proportions of white balls. The facts are given in Table I.

TABLE I

VARIABILITY OF TYPE PROPORTIONS FROM LINE TO LINE, WASHBURN COUNTY, WISCONSIN

(Unit: number of plots per line)

Cover type	Actual distribution σ	Lexian distribution σ_L	Bernoullian distribution σ_B
Aspen	11.2	12.7	6.1
Agricultural	21.5	22.1	5.4
Open brush	9.5	10.8	5.1
Pine	18.4	18.8	4.1
Oak	9.8	10.7	4.1
Northern hardwoods	14.2	14.6	3.7
Water	5.7	6.4	3.1
Spruce-tamarack	5.7	6.2	2.4
Hardwood swamp	4.1	4.6	2.0
Industrial	1.8	2.3	1.4

The total survey represents a Poisson sample.—Significant variability is thus shown among the several survey lines. These lines have tapped

different strata—parts of the total area in which the proportions of given cover types vary in more than a random or Bernoullian sense. It is as if the white balls had been more heavily concentrated in one part of the urn and some samples drawn from this part, others from other parts of the urn. Considering the entire survey of 3,456 plots as a single sample for the entire area, as it is in truth, it is a reasonable deduction from the above results that the complete sample has satisfied Poisson conditions, that is, that it contains representation of all the various strata for each cover type.[1] It is possible to break up the total sample in still a second way to test this conclusion. The plots are two miles apart at right angles to the survey lines, but along the lines they are only one-eighth mile distant. Plots at greater intervals along the lines could be expected on rational grounds[2] to give close agreement with the result of the total sample. To test this statement the total sample was broken up into sixteen sub-samples, as follows. The first sub-sample contained plots, numbers 1, 17, 33, 49, etc. on each line; the second sub-sample numbers 2, 18, 34, 50, etc., on each line. This process was continued to produce sixteen samples of 216 plots each, with the plots in each sample located two miles apart in the four cardinal directions. If the area is stratified and if these samples have approximately satisfied Poisson conditions the standard deviation of the sixteen should be less than the standard error of random sampling. If, however, the two-mile distance between plots allows some areas of a given cover type to be completely missed in some samples and included in others, this would have the effect of introducing Lexian conditions at once and the result might even be a standard deviation greater than the standard error of random sampling, since, as shown on page 260, Lexian influences exert a much more powerful effect upon the standard deviation than do Poisson influences. Table II supplies the data of this analysis.

The results, shown in this table are instructive despite the fact that they are based on only sixteen samples. In most instances, the actual standard deviation is less than Bernoullian expectation, which agrees with a priori predictions. The last column of the table tests the significance of deviation from Bernoullian expectation and shows generally non-significant deviation, but this conclusion is less tenable in view of the tendency for actual standard deviations to be always less than theoretical Bernoullian rather than to vary on either side of true Bernoullian. That they are more consistent generally with the Poisson standard errors seems of more significance and seems rationally to be

[1] Possible modifications of this statement with respect to cover types that occupy very small areas will be considered later.

[2] This point will be taken up later.

TABLE II

FOREST SURVEY AS A POISSON SAMPLE*—COMPARISON BETWEEN ACTUAL STANDARD DEVIATION OF 16 SAMPLES AND THEORETICAL STANDARD ERRORS**
(Washburn County, Wisconsin)

Cover type	Area per cent	Average number plots per sample of 216 plots	Standard deviation of actual distribution σ	Poisson standard error σ_P	Bernoullian standard error σ_B	$\frac{\sigma - \sigma_B}{\sigma_\sigma}$†
Aspen	26.10	56	6.55	6.40	6.46	+ .079
Agricultural	18.61	40	5.67	5.47	5.72	− .050
Open brush	16.26	35	6.27	5.38	5.43	+ .877
Pine	10.03	22	3.24	4.19	4.41	−1.49
Oak	9.85	21	3.53	4.32	4.38	−1.09
Northern hardwoods	7.93	17	3.00	3.82	3.97	−1.37
Water	5.13	11	2.05	3.21	3.24	−2.04
Spruce-tamarack	3.07	6.6	2.50	2.50	2.53	− .065
Hardwood swamp	2.14	4.6	1.26	2.10	2.13	−2.21
Industrial	.87	1.88	1.06	1.36	1.36	−1.11

*Each actual sample includes plots two miles apart in the four cardinal directions.

**The σ_p, used in formula for σ_P, is the standard deviation of the proportions on the eighteen lines, thus assuming that the survey lines represent the different strata.

† $\sigma_\sigma = \sqrt{\frac{1}{4N}\left(2npq + 1 - 6pq\right)}$ is the standard error of the standard deviation in a Bernoullian sample.

expected. The cases in which the actual sigma is greater than the Poisson sigma, may represent fluctuations of sampling or it may be a case where Lexian conditions of sampling have crept in, that is, where small areas of the given cover type were completely missed in some samples but included in others. If the latter be the true explanation it argues to the effect that two-mile spacing of plots is too great if Poisson conditions of sampling are to be maintained, and further to the effect that the errors assigned to such estimates should be Lexian errors.

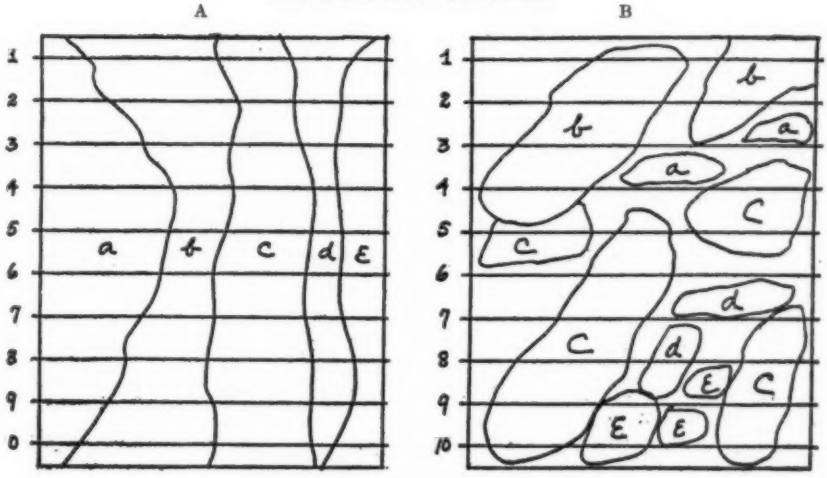
A rational view of land stratification.—A preliminary survey such as that of Washburn County is designed primarily to establish the procedure to be used in surveys of much larger areas. Because the preliminary survey is likely to be done on a more intensive scale than the larger survey it may err in being too accurate. That is, the accuracy realized when plots are located at one-eighth mile intervals on survey lines two miles apart may be unattainable when fewer plots to the mile are taken or when the survey lines are four, six, eight, ten or even twelve miles apart. But something approximating the latter distribution of plots may be required as a matter of sheer necessity because the money is not available for taking a more intensive survey. It may be a matter of getting what estimates are possible for the whole forest area of the United States with a given expenditure of money or getting nothing. Judgments based on the more intensive preliminary survey may need to be adjusted when the spacing of the plots and the distance between

the survey lines are greatly extended. This matter at least calls for attention and for whatever analysis is possible of the data of the preliminary survey.

The subject is best approached first on rational grounds. A tentative explanation was suggested for the peculiar results shown by the data of Table II, where successive samples of plots spaced at two miles in each direction showed in most instances a tendency to Poisson dispersion but where a few cases showed dispersion even greater than Bernoullian, suggesting the presence of Lexian conditions. How, it may be asked, could the latter intrude? A rational explanation is offered along the following lines. Given cover types such as forest, swamp or agricultural land may, if the survey area is crossed by important rivers, follow directions roughly parallel to the drainage systems—that is, their general contour may be that of long, narrow strips or bands; on the other hand, if the area is isolated at some distance from important rivers, it is possible that cover types are in more or less irregularly shaped plots, scattered somewhat promiscuously about over the area. The two contrasting situations, to use a simple figure, are like the stripes of a zebra and the spots of a leopard. Something approximating the cases pictured here are portrayed in the diagrams of Chart I, purely hypothetical in character.

CHART I

STRATIFICATION DIAGRAMS



The numbered lines are survey lines

One could well reason that in diagram A, if the horizontal dimension is fifty or a hundred miles a fairly good representation of the propor-

tionate area in each cover type could be made by plots along the lines at intervals for one, two or even four miles; and this could probably be done at a great saving over the casting of plots at intervals of one-eighth mile. As yet, remember, this is pure rationalization from the diagram, which deserves ultimately to be supported, if that be the outcome, by empirical analysis of data from the preliminary surveys. With fewer plots per line it would be less necessary to extend greatly the spacing between lines in order to keep within rigid expense limitations. But again with stratification, as in diagram A, lines at four-mile or even six-mile intervals would seem to give a very close approximation to the estimate based upon lines two miles apart.

On the other hand, consider diagram B. The survey lines on the diagram cut through many of the separate cover types but at the same time completely miss some of them. The larger areas seem certain of having proportionate representation in any sample that includes a set of lines covering the whole area at uniform intervals, but the small areas stand a chance of being entirely missed and the smaller they are the greater the chance. Correspondingly, the widely spaced plots on the survey lines can miss the smaller areas while being almost certain to represent all the larger ones. It seems highly probable that this sort of phenomenon is the explanation of the excessive dispersion of some of the distributions in Table II. Since each distribution in this table is composed of sixteen different samples, if some small areas are represented in a portion of the samples but not in the remainder the result would inevitably be Lexian dispersion.

The spacing of plots in a large survey.—If analysis of the survey data supports the above reasoning, it supplies an important judgment with respect to the method by which samples are to be taken on large areas. If funds available for the survey are sufficiently large the sample should be taken so as to have survey plots fall within all small areas that deserve a place in the final estimate; that is, the sample should assure Poisson conditions being observed. If money and time will not permit so detailed and intensive a survey then the ultimate error of sampling of the estimate should be calculated as a Lexian error, that is,

$$\sigma_L^2 = \frac{pq}{n} + \frac{n-1}{n}\sigma_p^2$$

and the σ_p could be calculated approximately as the standard deviation of the proportions on the several lines. If, on the other hand, the more intensive survey were possible, and Poisson conditions could be attained, the error of the estimate should be calculated as a Poisson error, viz.,

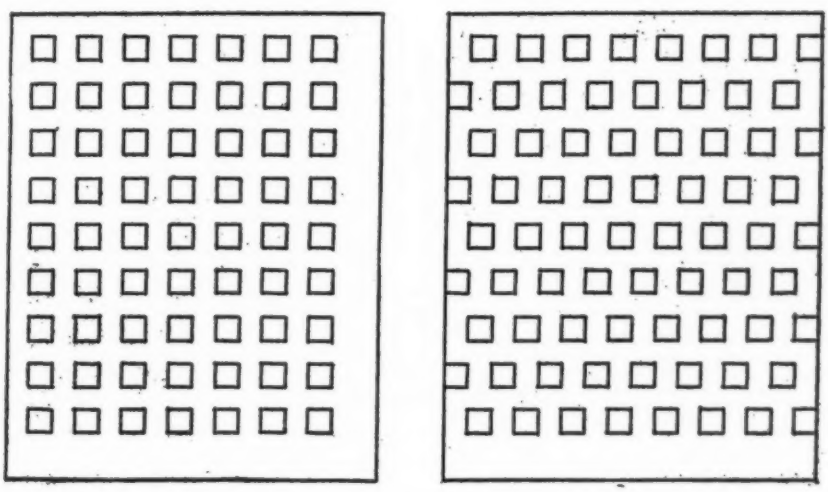
$$\sigma_{P^2}=\frac{pq}{n}-\frac{1}{n}\sigma_{p^2}.$$

The best representation of the total area in any case would seem to be so to arrange the spacing of survey plots on the lines and the distances between the lines that the entire area is covered uniformly by the plots in checkerboard fashion at equal distances in the four cardinal directions, as indicated in the diagrams of Chart II.

CHART II

SUGGESTED PLOT ARRANGEMENTS

A B



It would seem that the B arrangement would be preferred to A as a better means of sampling all parts of the area. With this general plan adhered to, the plots could be spaced just as closely as possible so as to keep the cost of the survey within budget limits.

As suggested above, the decision as to spacing between the survey plots is a purely budgetary one—space them as closely as the money available for the survey will allow. This decision having been made, it is then important to determine whether the sample taken is Poisson or Lexian. The preliminary surveys must be studied to settle the question at what distance between the survey plots the sample ceases to be Poisson and becomes Lexian. Table II gave part of the answer to this question for it showed that two-mile intervals between plots on each line resulted in Lexian dispersion in some cases. Further light may be shed on the problem by selecting as samples the plots on every other line, every third line, etc. In the first case there results two

samples each from lines four miles apart; in the second case three samples each from lines six miles apart, and so on. Despite the small number of samples, some indication of the effect of wider spacing of the lines is shown by the actual standard deviation of these samples and by comparison of these values with Bernoullian expectation from samples of the same size. These results are shown in Table III.

TABLE III

THE CHARACTER OF DISPERSION* OF PROPORTION OF AREA OF COVER TYPES IN RELATION TO DIFFERENT SPACING BETWEEN LINES

Cover type	Spacing between lines (in miles)				
	4	6	8	10	12
Northern hardwoods	2.14	1.74	1.93	2.22	2.23
Open brush	2.37	1.66	1.86	2.07	2.31
Agricultural	1.55	2.86	1.52	2.25	3.34
Oak	1.60	1.65	1.74	1.72	1.90
Spruce-tamarack	1.26	2.90	1.43	1.99	2.75
Aspen	1.75	1.46	1.54	1.40	2.45
Pine	.88	1.86	1.27	2.06	3.93
Water	.17	1.24	1.36	1.37	1.26
Industrial	.49	1.68	.73	1.37	1.28
Hardwood swamp	.17	.58	1.55	1.45	.80

* The values in the table are $\frac{\sigma_p}{\sqrt{\frac{pq}{n}}}$ where σ_p has been calculated as the standard deviation of the proportions in the actual number of samples available and $\sqrt{\frac{pq}{n}}$ is the Bernoullian expectation for a sample of n, using p obtained from the complete survey.

For the first six cover types listed in the table the results consistently show dispersion greater than Bernoullian for all given spacings between lines. The last four show ratios less than unity for four-mile spacings, and there are three instances of spacings greater than four miles giving ratios less than unity. These cases indicate the possibility that Poisson conditions of sampling have been attained in a few instances; but the burden of evidence given by the whole table, despite the meager character of the data, is to the effect that spacing of lines at intervals greater than two miles will most likely introduce Lexian conditions of sampling.

Another example of the effect of wide spacing of lines was obtained by sampling for jack pine, northern white cedar and swamps on available forest type maps of Bayfield, Sawyer, Vilas and Langlade counties, an area of 2,765,953 acres. Six lines per township with plots ten chains apart were used to represent the proportions of these types on the entire area. The total survey was then divided into six separate sub-samples, each with lines six miles apart, and each an estimate of the types on the

entire area. If six-mile intervals between survey lines introduce Lexian conditions into the sample it should be shown in the dispersion of the six proportions available for each cover type. Table IV shows the comparison of the actual standard deviations of the six samples with both Bernoullian and Lexian expectation where p_0 was the estimated proportions from the whole set of lines, both for individual counties and for the total area.

TABLE IV

DISPERSION OF SIX SAMPLES, EACH ESTIMATED FROM LINES SPACED AT SIX MILES, AND COMPARISON WITH BERNOULLIAN AND LEXIAN EXPECTATION

(Vilas, Sawyer, Bayfield and Langlade Counties, Wisconsin)

County	Number of townships	Standard errors (number of plots per sample)								
		Jack pine			Cedar			Swamp		
		Actual	Bernoulli	Lexis	Actual	Bernoulli	Lexis	Actual	Bernoulli	Lexis
Vilas	24	7.8	5.2	9.4	6.6	2.6	7.1	20.7	8.8	22.5
Sawyer	38	7.2	3.0	7.9	16.0	6.1	17.1	24.7	13.3	28.0
Bayfield	39	17.0	9.9	19.6	6.0	3.9	7.1	7.8*	9.3	12.1
Langlade	39½	2.0	1.3	2.6	3.3*	5.6	6.6	13.7	7.9	15.8
Total	120½	21.2	11.8	24.4	16.6	9.6	19.2	37.3	20.2	42.4

* Poisson dispersion.

The actual results shown in Table IV bear out the conclusions arrived at with respect to wide spacing of lines in Washburn County. The estimated type proportions for the large area arrange themselves in a Lexian manner instead of Poisson or Bernoullian.

One more body of evidence is available on this matter of wide spacing of survey lines and in this case the true areas are known and comparisons can, therefore, be made with them. Table V shows a comparison between the actual planimetered area of types in Menominee County, Michigan, and the areas estimated by both plot-count and linear distance methods. The comparison is made for lines spaced one mile apart and three miles apart. In every instance the estimate from one-mile spacing is closer to the planimetered area than the estimate from three-mile spacing. This is true for all the plot-count estimates and it is also true for the linear-measurement estimates. The differences between the two estimates, however, are generally small, although the percentage errors of the estimates vary more widely for the extremely small areas of pine and hemlock.

For this county it would seem a safe conclusion that either the one-mile or the three-mile survey lines give satisfactory estimates of the

true areas and that both have, to a sufficient degree, preserved Poisson conditions of sampling.

The total body of evidence now before us warrants the following tentative conclusions. The type of stratification of the total area to be surveyed should be considered carefully in laying out the method of taking the sample. Rational considerations indicate that spaces of six miles or more between survey lines are almost certain to introduce Lexian conditions of sampling whenever the stratification of the area is of the type shown in Diagram B of Chart I; and the empirical evidence supports the rational analysis. If large areas are to be surveyed where

TABLE V

COMPARISON OF AREA ESTIMATES BASED UPON ONE-MILE AND THREE-MILE SPACING OF LINES WITH PLANIMETERED AREA

(Menominee County, Michigan)

Cover type	Planimeter measurement	Per cent of total area			
		Estimated by plot-count method *		Estimated by measurement method	
		Lines one mile apart	Lines three miles apart	Lines one mile apart	Lines three miles apart
Aspen	27.60	27.28	26.00	27.81	29.85
Spruce, cedar, tamarack	22.20	22.37	23.70	22.35	24.19
Agricultural	18.86	17.50	16.83	17.83	16.75
Hardwood swamp	13.34	13.73	14.99	13.72	14.29
Northern hardwood	11.49	12.00	12.34	11.42	10.37
Wet land	2.93	3.50	3.69	3.42	3.74
Open brush	1.76	1.61	.74	1.73	.98
Oak	1.24	1.37	1.57	1.18	1.14
Pine	.34	.41	.43	.34	.18
Hemlock	.24	.23	.14	.20	.13

* Plots spaced 10 chains apart.

time and money limitations proscribe a survey on the intensive scale of the Washburn County project it would seem most sensible to locate the plots along the survey lines at wider intervals than ten chains. One plot per mile would be much more consistent with the necessarily wide spacing of the lines themselves. The survey lines should then be spaced at intervals of not more than two miles if that can be done within expenditure limits. This arrangement would give reasonable assurance of Poisson conditions of sampling for most cover types, and the error of the estimate should then be calculated by the Poisson formula.

If the above plot arrangement is still too expensive to be realized and if four- or six-mile intervals between survey lines must be used then the standard error of the estimate must be calculated from the Lexian formula.

ERRONEOUS MEASURES OF THE ERRORS OF AREA ESTIMATES

Schumacher and Bull [1] have recently considered at length the problems discussed in this paper. Their estimates of the sampling errors of the plot-count and linear-measurement methods of taking the sample are not supported by our analysis of the problem and it is believed that their conclusions are based on false reasoning.

They deal with an area of 503,000 acres in northeastern Louisiana, locating survey plots of one-quarter acre each at ten-chain intervals along lines three miles apart. There were 240 miles of line and 1918 plots in the total survey.

By plot count, there were 758 plots of cut-over restocking, or 39.52 per cent of total area for this cover type. The linear measurement estimate gave 7,654 chains of cut-over restocking out of a total of 19,169 chains of survey line, or 39.93 per cent.

Taking their analysis of the plot count estimate first, they calculate the Bernoullian error of sampling from the formula $\sigma_B = \sqrt{npq}$, which for a sample of 1918 plots gives $\sigma_B = 21.5$. "This means," they say correctly, "that, provided the conditions of random sampling have been fulfilled, it is to be expected that in two out of three additional sets of 1918 plots the number in the cut-over restocking conditions will be within 22 of the present count.

"But this is the point to be investigated. We need to know whether the standard deviation calculated in this way from the 1918 plots is identical with the standard deviation that would be obtained if all the quarter acres of land area in the two parishes could be thrown into an urn, thoroughly mixed, and a number of samples consisting of 1918 plots each could be drawn out." [2]

The answer to the query raised in this quotation is that $\sigma_B = 21.5$ is just the standard deviation that is to be expected from repeated samples of 1918 plots drawn from the hypothetical urn in accordance with the conditions of random sampling *but that it is not the standard deviation that is to be expected from repeated selections of 1918 plots taken as they were taken in the Louisiana survey.* For that survey most certainly did not, and should not, fulfill Bernoullian conditions. In all probability it fulfilled Poisson conditions, for the authors' description of the area indicates that it represents the "zebra" type of stratification. These circumstances are not recognized by Schumacher and Bull and they endeavor to test the adequacy of the error, $\sigma_B = 21.5$, by two methods: (1) mechanical grouping, and (2) survey line grouping. The former

[1] Determination of the errors of estimate of a forest survey, with special reference to the bottom-land hardwood forest region. *Journal of Agricultural Research*, Vol. 45, No. 12, December 15, 1932.

[2] Loc. cit., p. 744.

involves breaking up the total number of survey plots into twenty samples of about 96 plots each by taking for sample one plots numbers 1, 21, 41, 61, etc., for sample two plots 2, 22, 42, 62, etc., and similarly for the remainder of the samples. The procedure in setting up this distribution of samples is identical with our method of testing whether the total sample fulfilled Poisson conditions. Their survey line method of grouping involves taking as sub-samples the plots on each line. We used a similar arrangement of data to test stratification.

The authors are at a loss to explain why the standard deviation of twenty samples for method one is $\sigma=2.75$, whereas $\sigma_B=\sqrt{npq}$ for samples of 96 equals 4.79; and why the standard deviation of ten samples for method two is $\sigma=7.65$ whereas $\sigma_B=\sqrt{npq}=3.53$. Had they recognized that method one almost certainly fulfills Poisson conditions and method two with like certainty fulfills Lexian conditions, their perplexity over these results would have vanished. Neither of these two distributions has any logical connection with the conditions of Bernoullian or random sampling and neither standard deviation, therefore, can be expected to agree with $\sqrt{npq}$. They turn out to be just what the logic demands—the first less than Bernoullian, the second greater. The first gives the error of estimate of a single sample including every twentieth plot, when such single sample is used to represent the universe; the second, the error of estimate of a single sample composed of just one survey line, when such a sample is used to represent the universe. The first is a *good sample* of the total area, as is the author's total sample of 1918 plots; the second, an impossible one which no one would think of using.

Their test of the error of estimate by the linear measurement method proceeds from the same faulty foundation as that of the plot count method. They believe they can get a measure of the dispersion of a land area class by dividing up the whole length of survey line into segments, considering each segment a unit, making a frequency distribution of these units and calculating means and standard deviations. They construct three such distributions by breaking up the total length of line (1) into 480 half-mile units; (2) into eighty three-mile units; and (3) ten twenty-four mile units. The means and sigmas of these distributions are:

	Number Units	*Mean*	*Standard Deviation*
(1) Half-mile units.....	480	16	15.9
(2) Three-mile units.....	80	16	8.76
(3) Twenty-four-mile units.....	10	16	4.78

where both mean and sigma have in each case been expressed in "numbers of chains of cut-over restocking per half-mile." The standard deviations are then each divided by $\sqrt{n}$, that is, $n=480$, 80 and 10, to give standard errors of the mean: (1) .73; (2) .98; and (3) 1.51 and the authors conclude that "The increasing standard error of the means for units greater than one-half mile in length . . . clearly indicates that the variation in the data from the longer units, like the abnormal variation of the plot count when full line units were used, is correlated with 'place,'" and finally "that when areas in land units similar to these two Louisiana parishes are determined by the linear measurement method, the error of sampling of a land class may be deduced directly from the distribution of ungrouped one-half-mile units." [1]

The fundamental error in this procedure lies in the original conception upon which it is based—that the dispersion of half-mile, three-mile or twenty-four-mile segments into which the whole sample is divided has some connection with the dispersion of repeated samples, each of 240 miles of line spaced at three-mile intervals. The sigmas of their three distributions quoted above each has a perfectly definite meaning. Each gives the error of sampling of a single sample *composed of one line unit*, that is, one half-mile of line, or one three-mile line, or one twenty-four-mile line somewhere within the total area. These dispersions have no relationship to the standard deviation of repeated samples of 240 miles of line spaced at three-mile intervals throughout the area and the authors are in error in making any such use of them.

The authors follow this mistaken procedure to its logical conclusion, for they use it finally to make a choice between the plot count and the linear measurement methods of estimating land-class area. For this test they had true values obtained from planimetered areas on a soil type map. They divided the total area of the map into ten "line universes" extending across the map, each three miles wide. Survey lines were drawn through the middle of each universe and plots located at ten-chain intervals on each line. The number of plots of each soil type on each line were then recorded for the plot count estimate of each "universe" and the number of chains of each soil type on each line for the linear measurement estimate of each "universe."

The estimated proportion of area in each line universe was then compared with the true proportion for each soil type and an error of estimate calculated, using the formula $\sigma_B=\sqrt{\frac{pq}{n}}$ for the plot count data and $\frac{\sigma_{\bar{z}}}{n}=\frac{\sigma}{n\sqrt{n}}$ for the linear measurement data. Each estimate was ex-

[1] Loc. cit., p. 747 and 748.

pressed in units of this error, that is, in standard units. There were sixty-six of these ratios for the plot count and sixty-six for the linear measurement method, since nine of the "line universes" each contained seven soil types and the tenth one three soil types. Having obtained these sixty-six ratios expressed in standard units for each method, they then threw them into frequency distributions and calculated standard deviations. If the latter did not differ significantly from one, the authors concluded that the requirements of random sampling had been fulfilled, and the standard error of the estimate $\sigma_B = \sqrt{\frac{pq}{n}}$ could be used as a measure of sampling error.

The authors' calculated sigma for the distribution of plot count ratios was 1.72; for the linear measurement ratios 1.05 and they conclude that the error of estimate σ_B for the plot count method contains a bias and should be divided by 1.72, whereas the corresponding error for the linear measurement method is correct as it stands.

The difference between these two sigmas, 1.72 and 1.05, is explained on entirely different grounds. The standard error used to standardize the sixty-six values by the plot count method was $\sigma_B = \sqrt{\frac{pq}{n}}$, the Bernoullian sigma; that for the linear measurement method was obtained by their method, already described, of dividing each line universe into half-mile segments, constructing a frequency distribution of the number of plots for a given soil type per half-mile, then calculating $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ and finally dividing this value by n to get the standard deviation of a proportion. Had they used $\sigma_B = \sqrt{\frac{pq}{n}}$ to obtain this standard unit the differences between the two results would have vanished. The calculation can be made for the case of Sharkey clay plots for which they give the necessary data for line one. The proportion of clay plots is $p = \frac{170}{1160} = .14655$. Therefore, $\sigma_B = \sqrt{\frac{pq}{n}} = .01$ and the standardized value of the estimated proportion on line one is, therefore,

$$\frac{.132 - .146}{\sigma_B} = \frac{-.014}{.01} = -1.4\sigma$$

instead of being $+0.1\sigma$ as they have calculated it.[1]

[1] Their calculations are apparently in error since their distribution of plots on line one gives the proportion of Sharkey clay plots as .146 and not their figure, .129. The correction would make their value of the standardized ratio equal 0.28σ.

Rational considerations give a fairly strong guarantee that if a similar calculation were made for each of the sixty-six linear measurement ratios the result would be a distribution with a sigma of somewhere near their value of 1.72 for the plot count distribution. This is due to the very great probability that the soil types are not distributed on each line in a Bernoullian or random manner but rather in a Lexian manner, as has been shown at length in the earlier part of this paper.

The ultimate reason, however, for discarding their test of the two methods of estimate is that the dispersion of plots (or of chain lengths) along a given line has no logical connection with the dispersion of repeated samples made from the whole set of lines spaced at three-mile intervals over the entire area. The latter dispersion can only be estimated from the Poisson or the Lexian error formulas, as has been argued at length in the earlier pages of this paper.

ERRORS IN VOLUME

In the Wisconsin survey a complete census of all information wanted was taken on each survey plot. The discussion up to this point has dealt with only one item from this census, the cover type of the plot. Each plot was listed as belonging to one cover type and the estimated proportion of area for each cover type was taken as the ratio of the plots so listed to the total number of plots.

The volume of timber on each plot was also recorded. This figure itself is subject to an error of estimate since no method of counting trees and noting sizes suffices to give an exact measure of volume of standing timber, but for our purposes, the volume on each plot is taken as a known datum, not subject to error. The error that is thus neglected is probably of the second order of smalls as compared to the sampling error that we seek—the error in total volume estimated from the volume on the survey plots.

In addition, the size class and density of the forest were noted and the trees on the plot were tallied by diameter classes and species. Total height and age determinations were also made on a certain number of trees selected at random. Since the volume of forest stands varies from plot to plot with age, density and size of timber, the plots were classified according to these factors, or "forest conditions."

The process of estimating total volume of timber on the area must be built upon this information. Let it be approached as follows. Let

N = the *true* total area of the survey (in tenth-acre plots)

A = the *true* area covered by a given forest type (in tenth-acre plots)

V = the *true* total volume (in board feet) of timber of a given forest type

$a=\frac{A}{N}$ = the *true* proportion of total area in a given forest type

$v=\frac{V}{A}$ = the *true* average volume of timber per plot

$$\text{Therefore } N\cdot\ a\cdot v=N\cdot\frac{A}{N}\cdot\frac{V}{A}\equiv V.$$

Now, none of the above true quantities except N being known, suppose that estimates of a and v are obtained by a sampling process.

$$\text{Let } X_a=a(1+e_a)=\text{an estimate of } a$$

$$X_v=v(1+e_v)=\text{an estimate of } v$$

where e_a and e_v are relative errors in the estimates of a and v. Then

$$N\,X_a\,X_v=N\,a\,v\,(1+e_a)\,(1+e_v)=V(1+E)$$

$$\text{or}\qquad E=e_a+e_v+e_ae_v$$

$$=e_a+e_v \text{ approximately,}$$

since the e's being relative errors of estimate, their product is a second order of smalls that can be neglected in comparison with e_a or e_v. E is now the relative error in the estimated total volume of timber of a given forest type and is equal to the sum of the relative errors in the two components of volume, namely, proportion of total area and average volume per plot of the given forest type. E is, again, the relative error in a given sample and the standard error of sampling which is sought is the standard deviation of the frequency locus of the E's. Let

$$S=\text{standard deviation of } E$$

$$S_a=\text{standard deviation of } e_a$$

$$S_v=\text{standard deviation of } e_v.$$

Then $$S^2=S^2_a+S^2_v$$

since the variance of a sum equals the sum of the variances.[1] If it is now possible to calculate the values of S_a and S_v, we are in a position to make a final estimate of volume of timber and to state the standard error of sampling of this estimate.

The method of obtaining X_a, the estimated proportion of area with a given type of forest cover, has been considered at length and means of calculating the standard error of this estimate have been given. But this standard error has always been measured in absolute and not in relative terms. The relationship between an absolute and a relative standard error is given herewith:

[1] See Bowley, *Elements of Statistics*, pp. 288–9.

Let x_0 = any measurable quantity, subject to errors of measurement

x_1 = any estimate of x_0

$= x_0\,(1+e)$

σ_1 = the standard deviation of x_1

and $\sigma_1^2 = \dfrac{\Sigma x_1^2}{n} - \bar{x}^2_1$

where n values of x_1 represent n different samples.

Then $\sigma_1^2 = \dfrac{\Sigma[x_0(1+e)]^2}{n} - \left[\dfrac{\Sigma x_0(1+e)}{n}\right]^2$

$= x^2_0 \dfrac{\Sigma(1+e)^2}{n} - x^2_0\left[\dfrac{\Sigma(1+e)}{n}\right]^2$

$= x^2_0\left(1+2\bar{e}+\dfrac{\Sigma e^2}{n}-1-2\bar{e}-\bar{e}^2\right)$

$= x^2_0\left(\dfrac{\Sigma e^2}{n}-\bar{e}^2\right) = x^2_0\,\sigma_e^2$

$\therefore\ \sigma_e = \dfrac{\sigma_1}{x_0}$

In other words, a relative error of sampling is obtained by dividing the absolute error by the true value of the magnitude. The approximate value of this quantity which must be used in practice is obtained by dividing the absolute error by the best available estimate of x_0.

What then are the relative errors, S_a and S_v, to be used in the estimate of volume of timber? The standard error of sampling of the estimated proportion of total area covered by a given type of forest is σ_a where

$$\sigma_a^2 = \frac{pq}{n} - \frac{\sigma_p^2}{n} \text{ under Poisson conditions}$$

$$\text{or } \sigma_a^2 = \frac{pq}{n} + \frac{n-1}{n}\,\sigma_p^2 \text{ under Lexian conditions.}$$

The relative error is, therefore,

$$S_a = \frac{\sigma_a}{p_0}$$

where p_0 is the calculated proportion of area in the entire sample.

Assuming Poisson conditions the relative error, S_a, of aspen type would be:

$$n\,\sigma_a^2 = pq - \sigma_p^2$$

$$3451\,\sigma_a^2 = .192879 - .0034 = .189479$$

$$\sigma_a^2 = .0000549$$

$$\sigma_a = .00741$$

$$S_a = \frac{\sigma_a}{p_0} = \frac{.00741}{.261} = .0284$$

where σ_p has been calculated as the standard deviation of the proportions on the eighteen survey lines.

The average volume of timber per plot has yet to be calculated and its error estimated. Since volume of timber varies from plot to plot with age, density and size of timber, the plots were classified with respect to these factors for each forest type and average volume per plot in each such class was then calculated. Table VI shows these results for 931 plots of aspen in Washburn County. The average volume per plot for the entire aspen area is 42.05 cubic feet, as shown in the table. If the aspen area were a homogeneous universe the standard error of this estimate should be calculated by the formula for the sampling error of the mean, applicable thereto, namely,

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{47.31}{\sqrt{931}} = 1.55$$

but the universe is not homogeneous as was recognized in setting up the classification by forest condition in Table VI and as the results of this classification show, and this formula is therefore not a proper one to use.

TABLE VI

VOLUME ANALYSIS ASPEN TYPE, WASHBURN COUNTY, WISCONSIN

Age class years	Stocking t	Number of sample plots n_t	Average volume per plot (cu. ft.) $\bar{x}_t$	Standard deviation of plot volumes (cu. ft.) σ_t
1–20	Poor	178	16.91	19.22
	Medium	255	21.59	28.14
	Good	205	23.82	34.24
Totals		638	21.00	28.35
21–40	Poor	86	50.93	35.08
	Meduim	104	87.31	35.99
	Good	103	119.27	45.50
Totals		293	87.87	47.92
All classes		931	42.05	47.31

Of the total variance shown by volumes of the 931 plots, part is ascribable to the different kinds of forest condition, the remainder is here classed as errors of random sampling since no special causes of variation, other than forest condition, have been isolated. The problem is to separate the error of random sampling from the remainder of the variation. This may be done by calculating the error of a weighted average from Table VI, assuming the class averages, $\bar{x}_t$, and standard deviations, σ_t, approximately represent the true values. We define

$$\bar{x} = \frac{\Sigma n_t \bar{x}_t}{\Sigma n_t} = \text{average volume for universe}$$

$$= 42.05$$

$$\sigma^2_{\bar{x}t} = \frac{\sigma_t^2}{n_t} = \text{variance of } \bar{x}_t$$

$$\sigma^2_{\bar{x}} = \sum \left(\frac{n_t}{\Sigma n_t}\right)^2 \sigma^2_{\bar{x}t} = \sum \left(\frac{n_t}{\Sigma n_t}\right)^2 \frac{\sigma^2_t}{n_t}$$

$$= \frac{\Sigma n_t \sigma_t^2}{(\Sigma n_t)^2} = 1.10964$$

$$\sigma_{\bar{x}} = 1.0534$$

$$= \sigma_v$$

$$S_v = \frac{\sigma_v}{\bar{x}}$$

$$= .02505 = 2.505 \text{ per cent}$$

Table VII shows a similar estimate of average volume per plot of hardwood, with the accompanying error of the estimate, for Langlade County, Wisconsin.

This table contains estimates for both cord-wood and lumber. The calculations for the sampling error are given below for each.

	Cord-wood	*Lumber*
$\bar{x} = \frac{\Sigma n_t x_t}{\Sigma n_t} =$	.283	607.5
$\sigma^2_{\bar{x}} = \frac{\Sigma n_t \sigma_t^2}{(\Sigma n_t)^2} =$	.00003829	245.752
$\sigma_{\bar{x}} = \sigma_v =$	.00619	15.676
$S_v = \frac{\sigma_v}{\bar{x}} =$	.0219	.0258

TABLE VII

VOLUME ANALYSIS OF HARDWOOD TYPE, LANGLADE COUNTY, WISCONSIN

Size class of stands (inches)	Number of sample plots	Average volume per plot	Standard deviation of volume per plot
	Volume in cords, of trees 6–9 inches in diameter		
0–6	293	.082	.123
6–12	419	.454	.251
12+	334	.245	.182
Total	1046	.283	.252
	Volume in board feet, of trees 10 inches and over		
0–6	293	45	107
6–12	419	358	436
12+	334	1414	746
Total	1046	608	760

There is a tacit assumption in the above analysis of S_v which deserves to be recognized, for, while it is believed to accord with the facts in the case of the Washburn County survey, it might not do so in surveys of larger areas where plots are more sparsely distributed. The assumption is that in taking the sample to satisfy Poisson conditions with respect to area of the different cover types it also satisfies Poisson conditions with respect to the several classes of forest condition. That is, each class of forest condition is believed to be present in the whole universe in the proportions shown in the sample, subject only to an error of random sampling. But if a much larger area were sampled by a distribution of plots that would introduce Lexian conditions of sampling in respect to area of cover type it might also introduce Lexian conditions with respect to classes of forest condition. If such were the case the above procedure for calculating the error of sampling of the average volume per plot would not be a proper one. The latter should be calculated by use of a Lexian error formula. The problem would be to obtain a proper value for σ_p (or $\sigma_{\bar{x}}$ when applied to variables) in that formula. The values of $\sigma_{\bar{x}}$, given above for aspen and hardwood, would be the right values to use only in case the sample were taken wholly from one forest condition class, which of course is not done. It would seem most likely that a $\sigma_{\bar{x}}$ calculated from the average plot volume on each line of the survey would be more nearly in accordance with the character of the sample as taken. It seems impossible to resolve the issue from the data at hand, but it is just as well to recognize its existence.

All data and methods are finally in hand which are needed for making an estimate of total volume and for calculating its sampling error. Figures used in the illustration are for aspen forest in Washburn County, Wisconsin.

N = 553,687 acres—the total area of the county

= 5,536,870 plot units

$X_a = p_0 = .261$ = estimated proportion of area covered by aspen

$X_v = \bar{x} = 42.05$ cubic feet = estimated average volume of timber per plot

$S_a = .0284$ = relative error of sampling in X_a

$S_v = .02505$ = relative error of sampling in X_v

$N \cdot X_a \cdot X_v = 60{,}767{,}425$ cubic feet

= 60.77 million cubic feet = estimated total volume of aspen timber in the county

$S = \sqrt{S^2_a + S^2_v} = .03787$ = the standard error of sampling of the estimate,

which amounts to 2.3 million cubic feet, giving a final estimate, with error, of 60.77 ± 2.3 millions.

NET EMIGRATION FROM THE UNITED STATES TO CANADA: 1909–1931

By Norman Leon Gold

In the past forty years over two million people have left the United States with the intention of establishing residence in Canada. Comparatively few have remained. The contiguity of the two countries has facilitated easy movement, and in spite of emigration which, at its peak, exceeded 140,000 people in a single year, probably not over one-quarter of those who went to the Dominion took up permanent residence. No satisfactory study of the American [1] movement to and from Canada has been made. The object of this paper is to determine the net emigration from the United States to Canada, with the aid of statistics which have been made available since 1909.

For the movement into Canada, the official Canadian records are used.[2] These statistics were very inaccurate previous to 1908,[3] but with the establishment of adequate border inspection they became fairly satisfactory. The annual summaries have not included 1,709 United States citizens who entered Canada by ocean ports from 1901 to March 31, 1925, however. A more serious confusion is due to the fact that beginning April, 1924, returning Canadians were recorded separately. Although the Department of Immigration believed that such data had not been collected previously, it seems clear that from 1903 to June 30, 1914, at least, many thousand returning Canadians were listed in the reports, apparently through an oversight of the statistical department.[4] Consequently, the data have not been strictly homogeneous throughout the period considered, that is, from 1909 to 1931. However, they are the most accurate available and are used here to indicate the northward movement.

[1] In this paper an American is anyone going from the United States to Canada, or returning to the United States after living in Canada.

[2] These are contained in the annual *Reports of the Department of the Interior* up to 1918, and in the *Reports of the Department of Immigration and Colonization* after this time.

[3] See *Report of the Department of Immigration and Colonization for the Fiscal Year ended March 31, 1918*, Ottawa 1919, p. 6; and National Bureau of Economic Research, Inc., *International Migrations*, vol. i, New York, 1929, p. 358.

[4] See the materials supplied by W. D. Scott, former Superintendent of Immigration in *Canada and Its Provinces*, vol. vii, Toronto, 1914, p. 558; and in *Reports of the United States Commissioner-General of Immigration*, 1910, p. 142; 1912, p. 34; 1913, p. 170; and 1914, p. 212. According to these, 112,175 returning Canadians, presumably native born, or 12.39 per cent of the total emigration, went to Canada from 1903–04 to June 30, 1914.

See also *International Migrations*, vol. i; *op. cit.* pp. 357–359. The 313,822 Canadians who returned for permanent residence in Canada from April, 1924, to November 30, 1932, were not included in the total emigration, but were recorded separately.

The Canadian government has no record of the return movement of Americans, and no careful estimate has been made.[1] In 1929 the United States Department of State reported that there were 234,039 American citizens in Canada "whose residence . . . had a permanent or semi-permanent character," but this figure left out the large number of people who have renounced their citizenship.[2] Another source for the net movement of Americans is the Canadian census of United States-born persons living in the Dominion. However, the emigration to Canada included naturalized citizens and aliens, as well as native-born Americans.[3] Moreover, neither of these sources makes possible a reliable estimate of the *yearly* movement to and from Canada.

There is one other source for the return movement of Americans. The annual *Reports of the United States Commissioner-General of Immigration* have listed the return of United States citizens. For 1909 the figure was an approximation; the subsequent data included United States citizens returning through Atlantic and Pacific ports of Canada, as well as through the border stations, after an absence of a year or more. The most serious error in these figures, as a record of the return movement from Canada, lies in the fact that many persons enumerated arrived from over-sea residences, and merely entered the United States through Canadian stations.

The materials compiled from the Canadian and American sources are summarized in Table I.

It is certain that many emigrants from the United States to Canada were not United States citizens, and it may be assumed that the return movement also included many aliens. It follows that a calculation of the net movement, based on figures for returning United States citizens alone, would be incomplete. The problem is: if the United States figures for withdrawals of American citizens are accepted, what percentage of the total return movement do they represent?

It is impossible to find from the Canadian reports what proportion of the total movement to the Dominion was United States citizens. However, the United States Bureau of Immigration has classified emigrants to Canada, to indicate whether they are citizens or aliens, since 1910. Since it is more difficult to obtain a complete enumeration of emigrating than of immigrating people, the materials are probably

[1] Mr. Roland Wilson prepared an estimate of the net movement to Canada from all countries for the period 1868 to 1925 (*Canadian Historical Review*, vol. xiii, no. 2, June, 1932, pp. 157–183). For the most part, his study is based on computations from the decennial censuses. In the writer's opinion, many serious errors are present in that study. Though Mr. Wilson used sources similar to those in the present paper, the details of manipulation and the final results are quite different.

[2] Walter F. Willcox in National Bureau of Economic Research, Inc., *International Migrations*, vol. ii, New York, 1931, pp. 105–107.

[3] Cf. Dominion Bureau of Statistics, *Sixth Census of Canada, 1921*, vol. ii, p. xiii.

TABLE I

EMIGRATION FROM THE UNITED STATES TO CANADA, AND THE RETURN OF UNITED STATES CITIZENS: 1909–1931*

Year**	Emigration from U. S.†	Return movement of U. S. citizens	Year	Emigration from U. S.	Return movement of U. S. citizens
1909	72,349	15,000	1923	19,955	30,133
1910	116,374	28,831	1924	19,580	23,219
1911	119,753	39,955	1925	15,693	45,264
1912	143,251	45,051	1926	19,718	20,969
1913	130,664	60,368	1927	22,441	21,102
1914	90,818	50,482	1928	26,687	22,543
1915	42,529	56,640	1929	32,507	22,182
1916	40,328	33,709	1930	28,947	9,134
1917	78,676	46,631	1931	19,577	19,478
1918	53,741	29,096			
1919	44,003	25,810	1912–1921	716,127	389,709
1920	48,624	19,712	1922–1931	231,044	235,023
1921	43,493	22,210			
1922	25,939	20,999	1909–1931	1,255,647	708,518

*Compiled from Canadian reports on Immigration, and from Table I of the annual *Reports of the United States Commissioner-General of Immigration.*
**Fiscal year ended June 30. †Recomputed for a fiscal year ended June 30.

not so accurate as those compiled by the Canadian government.[1] Nevertheless, the two series give approximately the same results, the Canadian total for the 22-year period, 1910–1931, being 1,183,298, and the United States total, 1,234,042. A computation of the percentage of naturalized citizens in the total emigration is presented in Table II.

TABLE II

PER CENT OF UNITED STATES CITIZENS IN THE TOTAL EMIGRATION FROM THE UNITED STATES TO CANADA: 1910–1931*

Year**	Per cent	Year	Per cent
1910	67.46	1923	74.18
1911	64.08	1924	76.48
1912	67.18	1925	71.98
1913	64.53	1926	84.23
1914	66.18	1927	88.15
1915	55.48	1928	87.86
1916	59.25	1929	88.54
1917	70.17	1930	86.86
1918	55.36	1931	82.94
1919	66.25		
1920	70.13	1910–1922	64.82
1921	66.01	1923–1931	83.98
1922	69.98		

*Computed from *Annual Reports of the United States Commissioner-General of Immigration.*
**Fiscal year ended June 30.

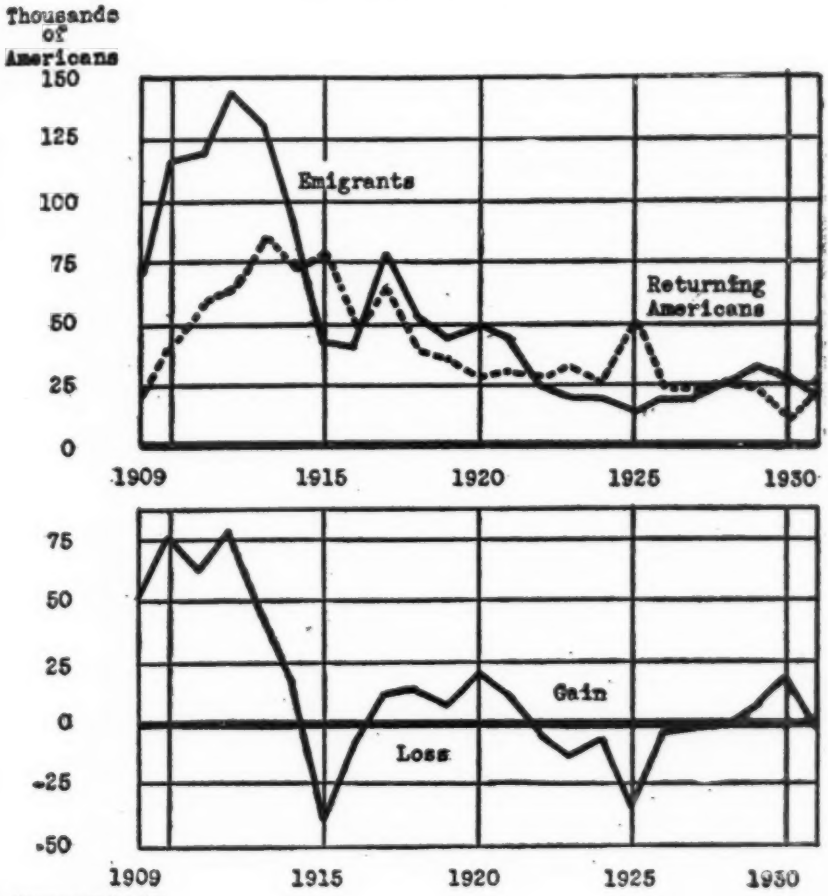
From 1910 to 1922 inclusive, United States citizens were 64.82 per cent of all who went to Canada from the Republic. By 1922, the quota law in the United States, and the more restrictive Canadian

[1] The American figures have consistently exceeded the Canadian figures from 1926 to 1931. This is due partially to the separate compilation, made during this time by the Canadian government, of returning Canadians.

immigration regulations, deterred many aliens from leaving the United States. During the nine year period 1923–1931, the citizens were 83.98 per cent of the total movement.

It might be assumed from this that to 1923 the United States figures for the return of citizens were equal to 64.82 per cent of the total with-

EMIGRATION OF AMERICANS FROM THE UNITED STATES TO CANADA AND THE RETURN MOVEMENT: 1909–1931*



*Source: Table III.

drawals, and that after this time they were 83.98 per cent of the return movement. However, the emigration of United States citizens probably included a higher proportion of transients than did the movement of aliens. The citizens frequently left their families and friends, announced to the Canadian immigration officials that they were entering

Canada for permanent residence, but easily and quickly returned to their former homes if opportunities were unsatisfactory. Often the aliens did not have the same family ties. When they left they made a definite break, and indeed, especially after 1921, they were likely to have difficulty in returning to the United States. It is assumed therefore, that the return of Americans included a higher ratio of United States citizens than did the emigration. On this assumption an estimate is made that citizens comprised 70 per cent of the total return movement from 1909 to 1923, and 90 per cent of the total after 1922.

TABLE III

NET EMIGRATION FROM THE UNITED STATES TO CANADA: 1909–1931

Year*	1 Emigration from U. S.**	2 Return to U. S. (computed)†	3 Net annual emigration 1–2	4 Total emigration (cumulative)	5 Net emigration (cumulative)	6 Ratio of net to total emigration (cumulative) 5÷4
1909	72,349	21,429	50,920	72,349	50,920	70.38
1910	116,374	41,187	75,187	188,723	126,107	66.82
1911	119,753	57,079	62,674	308,476	188,781	61.20
1912	143,251	64,359	78,892	451,727	267,673	59.26
1913	130,664	86,240	44,424	582,391	312,097	53.59
1914	90,818	72,117	18,701	673,209	330,798	49.14
1915	42,529	80,914	−38,385	715,738	292,413	40.85
1916	40,328	48,156	− 7,828	756,066	284,585	37.64
1917	78,676	66,616	12,060	834,742	296,645	35.54
1918	53,741	41,566	12,175	888,483	308,820	34.76
1919	44,003	36,871	7,132	932,486	315,952	33.88
1920	48,624	28,160	20,464	981,110	336,416	34.29
1921	43,493	31,729	11,764	1,024,603	348,180	33.98
1922	25,939	29,999	− 4,060	1,050,542	344,120	32.76
1923	19,955	33,481	−13,526	1,070,497	330,594	30.88
1924	19,580	25,799	− 6,219	1,090,077	324,375	29.74
1925	15,693	50,293	−34,600	1,105,770	289,775	26.21
1926	19,718	23,299	− 3,581	1,125,488	286,194	25.43
1927	22,441	23,447	− 1,006	1,147,929	285,188	24.84
1928	26,687	25,048	1,639	1,174,616	286,827	24.42
1929	32,507	24,647	7,860	1,207,123	294,687	24.41
1930	28,947	10,149	18,798	1,236,070	313,485	25.36
1931	19,577	21,642	− 2,065	1,255,647	311,420	24.80
1912–1921	716,127	556,728	159,399			
1922–1923	231,044	267,804	−36,760			
Total 1909–1931	1,255,647	944,227	311,420			

* Fiscal year ended June 30.

** Recomputed from the annual *Reports of the Department of the Interior* (Canada) for the fiscal years 1909–1917 and the annual *Reports of the Department of Immigration and Colonization* for the fiscal years 1918–1932.

† Based on Table I of the *Annual Reports of the United States Commissioner-General of Immigration* for the fiscal years 1909–1931. (See Table I above.) The figures for returning United States citizens were taken as 70 per cent of the entire movement for the years 1909–1922, and 90 per cent of the movement for the years 1923–1931.

From these approximations the return movement to the United States is calculated. Subtraction of the annual figures thus obtained, from the Canadian figures on emigration, gives the net yearly movement. In order to obtain a figure for the total net emigration for the 23 years, the annual figures are cumulated.[1]

[1] The net figures do not include any deductions for deaths.

In spite of the large amount of conjecture involved in the calculation, the results seem fairly reliable. They may be checked roughly by the census records of the increase or decrease in the United States-born residents in Canada. According to the table, 716,127 emigrants went to Canada from the United States between the fiscal years 1912 and 1921. Since it was determined that 64.82 per cent of the total were United States citizens, 464,194 may be so included. During the same time, 556,728 Americans returned to the Republic, of whom 389,709 were United States citizens. This meant an increase of 74,485 citizens in the ten years. The Canadian census reported the increase of the United States-born to be 70,344 from 1911 to 1921.

The United States-born population decreased 29,450 from 1921 to 1931. According to Table III, 231,044 Americans emigrated to Canada in the ten years ended June 30, 1931. Of these 189,061 were assumed to be United States citizens.[1] During the same period 267,804 persons left Canada to return to their former domicile, of whom 235,023 were citizens.[2] This indicates a decrease of 45,962 citizens.[3]

In the 23 years from 1909 to 1931, Canada increased its population by approximately 311,000 Americans or 24.8 per cent of all who left the United States for Canada—a fact which should have considerable bearing on Canada's future immigration program. Except in 1915 and 1916 there were steady though declining gains to 1922. The return movement was most serious after 1921. In practically every year Canada lost many thousand persons more than it was able to attract, or would permit to enter the country.

[1] This is obtained by taking 64.82 per cent of the 1922 emigration, and 83.98 per cent of the 1923–31 emigration, i.e. (25,939×.6482) + (205,105×.8398) = 189,061.

[2] This figure may be obtained directly from Table I above, or by taking 70 per cent of the return movement for 1922, and 90 per cent of the subsequent return, i.e. (29,299 ×.70) + (237,803×.90) = 235,023.

[3] 16,512 more than the decline reported by the census.

RESEARCH AND STATISTICAL PROGRAM OF THE FEDERAL EMERGENCY RELIEF ADMINISTRATION

By Emerson Ross

The Division of Research and Statistics in the Federal Emergency Relief Administration was created shortly after the organization of the Administration in May, 1933. The activities of the Division since that time in the direction of collecting original data not previously available fall into three main categories. The first major job undertaken was the establishment of a regular monthly reporting system to show the volume and trends of unemployment relief in the United States in terms of both families and dollars. This reporting system was later extended to cover the Civil Works program while this was in operation. While these reports were successful in showing in a quantitative manner the basic unemployment relief statistics, additional information was necessary for a more thorough understanding of the relief problem. In the second category, then, would fall the research program involving a more detailed examination by the questionnaire schedule method of particular aspects of unemployment relief. The third phase, that of correlating and advising on statistical studies financed by relief and Civil Works funds but not directly under the supervision of the Relief Administration, began during the period when Civil Works funds were granted to other agencies for the purpose of gathering statistical information as a means of giving employment to white-collar workers. After the cessation of the Civil Works program, state and local relief administrations were required to send an outline of all projected statistical studies to the Federal Emergency Relief Administration for recommendation and approval.

In addition to the collection of new material, the Division of Research and Statistics has made use of available material wherever possible. Investigations have been made and a current record is kept of the financial conditions of state governments; national and local economic conditions are interpreted and analyzed in terms of unemployment relief conditions; and records are maintained of all known studies pertaining to unemployment relief.

MONTHLY UNEMPLOYMENT RELIEF STATISTICS

The Federal Emergency Relief Administration requires each state relief administration to submit monthly reports of unemployment relief statistics for each county and most of the large cities in the state. The state emergency relief administrations in turn obtain the data from emer-

gency relief officials in the counties or other local political subdivisions. About 2,800 out of the slightly more than 3,000 counties in the country are covered by these reports, representing all counties in which unemployment relief is granted from public funds. To establish this reporting system on an adequate basis, it has been necessary to organize or to strengthen the statistical services in many of the states and to expend a substantial amount of effort in improving local records and accounting and reporting procedures. A manual has been prepared outlining in detail an accounting procedure which is mandatory for those states under the direct supervision of the Federal Emergency Relief Administration. By the use of some of the procedures outlined in this manual, the derivation of reliable statistics is simplified.

At the present time these monthly reports yield data by counties showing the number of families receiving direct relief, the number of families receiving work program earnings, and the net unduplicated number of families receiving either or both types of relief. The figures show for each type of relief, the number of single resident persons, the number of two or more person families, and the number of persons in the families. The dollar figures of relief granted are also shown and correspond to the above classifications. From these basic data, it is possible to compute the proportion of families receiving relief in any area to the total number of families in that area as reported by the 1930 census, the proportion of the total population receiving relief by a similar computation, and the average amount of relief benefits per family.

A new report, still partially on an experimental basis, was instituted during February, 1934, and called for the number of cases opened or added to the relief rolls during the month and the number of cases closed during the month. The number of cases opened was separated into those receiving public unemployment relief for the first time and into those who had received public unemployment relief at some previous time and were being reopened. For each of the above categories the number "opened" or "closed" was further divided into "due to Civil Works employment" and "for all other reasons."

In addition to the above data, the monthly reports give the sources of funds obligated as against Federal, state or local funds, the amount spent for administration, and the amount spent and the persons benefited by special programs such as relief to transients and the emergency education program.

In all instances, the statistics are limited to relief given from public funds only; that is funds provided by municipal and county governments, by state governments, and by the Federal government.

In connection with the new work program, which went into effect in April, 1934, the number of persons working, the number of employee hours worked, and the amount of earnings are reported weekly. The quantities, types, and value of material used on work projects and the sources of the funds used are to be reported monthly.

When the Civil Works program was initiated, the various state and local relief agencies were designated as Civil Works administrations, and with the relief organizations as a nucleus it was possible almost overnight to develop an effective disbursing, accounting, and reporting organization. Reports on employment, employee hours, and payrolls were received weekly throughout the program. Since copies of all payrolls were sent to Washington, it was possible to make sample studies for selected weeks showing a frequency distribution of the number of employees by wage rate, hours worked, and weekly earnings class intervals. As reports of completed projects are received, statistics by types of projects and the amounts spent for various kinds of materials will be available.

The various types of monthly statistical information concerning unemployment relief described above became available at different times as the reporting system was expanded. Total obligations incurred, broken down by sources of funds, are available for each state since January, 1933. In approximately one-half the states, the number of families receiving relief is also available since January, 1933. Beginning with April, 1933, the number of families and amount of funds expended are available by counties for all states. In July, 1933, another expansion took place and the number of single resident persons was shown separately from 2 or more person families and the total number of persons receiving relief was reported. The reports concerning cases opened and closed were inaugurated in February, 1934. Although the earliest forms provided for reports on work relief, it is not believed that these reports became reliable until the month of April, 1934, after the cessation of the Civil Works program. Civil Works statistics, of course, cover the period from the beginning of the program in the middle of November, 1933, to its cessation.

RESEARCH PROGRAM

A census of resident families and persons receiving unemployment relief in October, 1933, was the first of a series of research studies designed to provide more detailed facts concerning unemployment relief than could be obtained on a monthly basis. The October count of individuals by age, color, and sex, and of families by size and color has been followed by studies of rural relief families and their non-relief neighbors, studies of

families coming on or leaving the active relief lists, studies of transients receiving relief, surveys of the occupational characteristics of relief persons, studies of former members of the Civilian Conservation Corps, and a number of inquiries into unusual local situations. With the exception of the October census and the Civilian Conservation Corps studies, which covered the whole United States, these surveys have been limited to representative sample groups of cities and rural counties. For the most part the studies were begun under a Civil Works Administration grant, as a Federal statistical project. After the Civil Works program ended, in April, 1934, the survey staffs have been drawn principally from the local work divisions of the emergency relief administrations. Part of the central executive and technical personnel has been provided by a grant of $50,000 from the Rockefeller Foundation, administered through the Social Science Research Council.

The October, 1933, unemployment relief census answered for the first time a number of questions that had long been asked by public officials and other persons concerned about relief. The census covered practically every family receiving unemployment relief from public funds in the United States in October, 1933. Information obtained by this census was restricted to data available on case card records in relief offices throughout the country. Information was secured showing for each family the place of residence, size of family, the age, sex and race of persons in the family and the relationship of family members to the head. The first report of the results of this census has been published showing, by geographic divisions, by states, and for the 93 cities with a 1930 census population of 100,000 or more, the number and size of families receiving relief and the age, sex, and race of individuals comprising these families. A second report, breaking down the above data as between urban and rural areas for geographic divisions and states, is now (June 28, 1934) in press. This second report also shows the number of families and persons receiving relief classified by rural and urban areas and by race for each county in the country. The number of relief families and persons classified by race is shown for each city with a population over 25,000. A third report, an analysis of family composition, is to follow.

Following the October census, special studies were launched in representative cities and rural counties. In 52 counties within 20 states the economic and employment status, occupational history, composition of household, and relief history of rural relief families were studied in comparison with their non-relief neighbors. Field work on this study had been completed by June 16, 1934, with 21,000 schedules filled and the work of editing, coding, and tabulation well advanced.

The census was also supplemented by sample studies in 50 rural counties and 39 cities of families for whom relief was initiated, resumed, or discontinued during the period from November, 1933, to April, 1934, inclusive. These studies were designed to determine the effect of private and governmental employment in increasing or decreasing the relief load and to compare the occupations and characteristics of families leaving the relief rolls with those which were being added to them. On June 16, 1934, schedules had been secured for 93,000 rural families and 493,000 urban families and the work of editing, coding, and tabulating was nearing completion.

Transients and non-residents were not included in the October census but were studied in thirteen cities selected as strategic points in lines of migration. Between April 1 and June 16, 1934, schedules for 68,000 transients were secured. This study was aimed to secure the employment history and migration experience of transient persons and families, to throw light on the causes of migrancy and to provide data concerning a number of other factors connected with transiency.

In June, 1934, a study of occupational characteristics of persons receiving relief was started. This study was designed to record the age, sex, color, nativity, training, usual occupation and industry, present occupation if employed, period of unemployment, disabilities, and other personal and occupational characteristics of workers and potential workers in relief families. This study covered a representative group of 79 cities. Four regional offices were established to decentralize and speed up the tabulation. Not only the major personal and occupational features of workers on relief, but also the composition of their families will be analyzed. The study will show the number of families which have no readily employable persons, and the number of workers in families including employable persons. It will picture the kind and amount of employment necessary to free families from dependence on relief.

A survey has also been made of all Civilian Conservation Corps enrollees who left camp during or at the conclusion of the first six-months term, to determine why they left and the number who have since secured employment. A second survey, securing the same information for all enrollees who left during or at the end of the second term of camp, was begun in June, 1934.

In addition to these larger studies certain special investigations have been carried on for administrative purposes. Studies have been made of placer gold mining, copper and coal mining in selected areas, tung oil production, typical stranded communities, high relief counties, and several problem areas. In June a special staff was organized to make

quick reconnaissance surveys of rural problem areas. These special investigations will provide information concerning the families, resources, and administration of relief in areas, such as parts of the drought region, in which distress has been particularly acute and lasting. They have been designed primarily for immediate administrative guidance. The research projects described in the preceding paragraphs provide more precise and detailed facts and analyses designed to permit effective planning and administration of unemployment relief over a longer period.

Within less than a year a large research organization has been formed and great quantities of data have been collected. Tabulations of these data have proceeded according to planned schedules and are being analyzed. Some of the results have been issued in preliminary form and a number of the final reports will go to press during the summer, while additional studies go into the field to take the place of those completed.

CIVIL WORKS STATISTICAL STUDIES

Among the important statistical results which have come from the efforts of the Civil Works Administration, none is more striking than the bodies of data that have been gathered throwing light on various aspects of the emergency. A large number of statistical and survey projects were undertaken by local and state authorities. A complete population census was taken in Chicago, Illinois; a widespread unemployment survey was made in Massachusetts; local real property inventories were undertaken in many centers. Most of these, however, were of local interest.

By far the most significant studies were those carried out by the various agencies of the Federal Government, financed by Civil Works grants and supervised by the Federal Civil Works Administration and the Central Statistical Board. As a result, the country now has reliable information on such topics as the extent and trends of tax delinquency, the condition of farm homes and equipment, the status of urban and rural home financing, rural as well as urban commodity prices, and on a variety of other subjects. These Federal projects constituted a major source of work for unemployed statisticians and statistical workers throughout the country, as well as for many other groups of white-collar workers.

In all, twenty-four projects were undertaken. Original estimates of cost totalled twelve million dollars and the amount actually expended was slightly over ten million dollars. Their cost ranged from about one thousand dollars to two and a half million dollars, practically all being

expended for labor. Their personnel varied from a few clerks in the District of Columbia to ten thousand enumerators, clerks, supervisors and other personnel in every part of the nation. The number of persons employed on these projects reached nearly forty thousand during the maximum week.

The ten most important projects in terms of subject matter were: Unemployment Relief Census and Supporting Studies (mentioned in the previous section of this report), Real Property Inventory, Urban Tax Delinquency, Rural Tax Delinquency, Farm Housing Survey, Census of American Business, Railway Employees Study, Cotton Statistics, Compilation of Meteorological Data, and a Survey to Determine Possibilities of Employment for the Deaf and Hard of Hearing.

Next to these in significance because they supplement and develop existing important series were the following seven: Retail Price Study, Prices Farmers Pay, Employment and Payrolls, Trial Census of Unemployment, Compilation and Analysis of Statistics of Employment Offices, Census Record Preservation, and Photo Mapping.

Four projects were primarily designed to make existing material more available and were small-scale but worthy: Consumption Statistics Indexes, Drafting Work, Historical Series on Foreign Trade, and Russian Rainfall Studies.

Two studies were solely of local interest: Farm Land Use Survey (South Carolina), and Retabulation of 1930 Census Records (New York State).

One project, the Study of Subsistence Homesteads Records, changed its function from that of gathering material for general statistics to the collection of data for administrative purposes only.

Many of these projects have been described in earlier issues of our JOURNAL and others will be in the future.

FERTILITY OF FAMILIES ON RELIEF [1]

By Samuel A. Stouffer, *University of Wisconsin*

Do families after going on relief conceive more children than comparable families not on relief? This question has both scientific and popular interest.

The present study is submitted as a minor contribution, to be matched with other studies, toward a factual answer to this question. The findings apply only to Milwaukee, Wisconsin; and the suburbs of Shorewood, Wauwatosa, and West Allis. Extensive further research would be needed to reveal how far the findings in the Milwaukee metropolitan area apply elsewhere.

Our data show that there were 35 per cent more confinements between October 1, 1930, and December 31, 1933, among 5,520 families on public relief than among a control group of non-relief families. However, this is apparently an understatement of the difference. Calculating confinement rates per 1,000 months of exposure to the risk of pregnancy, we find that the relief group has an excess of 43 per cent. The basic data are exhibited in Table I.

These data include only confinements taking place at least nine months after a family went on relief, or within a comparable period among the control group. Obviously, a comparison based on the total number of children ever born to relief and non-relief families would be different, since families with children are more likely to need relief than childless families in the same occupational group.

The 5,520 relief families represented in this study comprise all the families applying for relief between January 1, 1930, and March 31, 1933, among a group of over 25,000 Milwaukee families studied as part of a larger investigation, nearing completion, of the accuracy of birth registration and differential fertility in Wisconsin cities. The entire group of Milwaukee families represents about 97 per cent of the couples now known to reside in Milwaukee and suburbs who were married between 1919 and 1930, inclusive, in any Wisconsin city of over twenty thousand or certain cities in adjacent states;[2] exclusive of cases where:

[1] Acknowledgment is made of the valuable assistance of L. W. Hutchcroft, Chief Statistician of the Wisconsin Board of Health, Malcolm Morrow, Mae G. Bergfors, and Paul C. Glick. The study was financed largely by Federal Emergency Relief, with supplementary funds from the University of Wisconsin, and the Wisconsin State Board of Health.

[2] These cities were Waukegan, Belvidere, Woodstock, Rockford, and Freeport, Ill.; Winona and Duluth, Minn.; and Menominee, Mich. The 3 per cent not included in the study failed to respond to mail inquiries. A sample check on these families, made by inquiring of relatives and acquaintances, revealed no marked differences as compared with the 97 per cent who responded to mail inquiries. In any event, the proportion is not large enough to bias the results materially.

TABLE I

MILWAUKEE, WISCONSIN—CONFINEMENTS AMONG 5,520 RELIEF FAMILIES OCCURRING LATER THAN 9 MONTHS FROM FIRST APPLICATION FOR PUBLIC RELIEF, AS COMPARED WITH CONFINEMENTS AMONG "CONTROL GROUP" OF 5,520 NON-RELIEF FAMILIES

Data include births from October 1, 1930, to December 31, 1933

Occupational classification of husband and probable religion as indicated by officiator at marriage	Relief families				Non-relief families				Percentage by which number of confinements among relief families exceeded that of non-relief families	Percentage by which confinement rate of relief families exceeded that of non-relief families
	Number of families	Aggregate months of exposure to risk of pregnancy	Number of confinements	Number of confinements per 1,000 months exposure	Number of families	Aggregate months of exposure to risk of pregnancy	Number of confinements	Number of confinements per 1,000 months exposure		
Clerical workers										
"Catholic"	206	4,220	39	9.2	206	4,474	27	6.0	44	53
"Non-Catholic"	362	7,335	61	8.3	362	7,823	46	5.9	33	41
Skilled workers										
"Catholic"	1,013	22,623	212	9.4	1,013	23,030	194	8.4	9	12
"Non-Catholic"	1,261	27,700	246	8.9	1,261	29,260	164	5.6	50	50
Semi-skilled and unskilled workers										
"Catholic"	1,103	25,615	282	11.0	1,103	26,644	225	8.4	25	31
"Non-Catholic"	1,575	35,057	382	10.9	1,575	37,733	248	6.6	54	65
Total	5,520	122,550	1,222	10.0	5,520	128,964	904	7.0	35	43

(a) the bride was over 25 and the bridegroom was over 30 at marriage; (b) there had been a second marriage of either party; and (c) the bridegroom was not a resident of some Wisconsin city of over twenty thousand at time of marriage. A simpler method of choosing the group would have been adopted had the study here reported been the sole objective. As it was, the present study merely operated on data already collected for the larger investigation.

The following steps were taken as part of the larger study:

1. The names of a couple qualifying for inclusion were copied on an index card from the marriage license record. The latest address of the couple was obtained from city directories, of which a complete file for all Wisconsin cities was available. At the same time, the husband's occupation reported in the directory was recorded.

2. The occupation was coded for inclusion in one of four groups: (1) professional or proprietary, (2) clerical, (3) skilled worker, (4) semi-skilled or unskilled worker. The classification scheme followed substantially that of W. F. Ogburn and Clark Tibbitts in *Recent Social Trends.*[1]

3. The name of the officiating clergyman or magistrate was checked against a practically complete file which had been assembled, and the religion of the clergyman (Catholic, Protestant, or Jewish) entered on the card.

4. A schedule was mailed to the couple requesting a list of the names of children born alive and now living, born alive and now dead, and stillborn, with date and place of birth. Request also was made for information as to date of death of husband or wife if a death had occurred and approximate date of separation if separation or divorce had occurred. One object of the inquiry was to help parents check the registration of their children's births; the response, consequently, was excellent. About 60 per cent of all married couples located by the post office replied to the first letter and the remaining 37 per cent replied to second, third, fourth, or fifth letters.

5. The place and date of birth of the children were recorded on the index cards.

The subsequent steps in connection with the present study were as follows:

6. All cards of Milwaukee county residents were checked against the indexed files of the Milwaukee County Department of Outdoor Relief. The date of first application for relief was recorded on the card. Since our card and the Milwaukee record had the maiden name of the wife there was a practically certain check against error. Although some relief cases may have been missed because of different spelling of names,

[1] *Recent Social Trends in the United States*, Vol. I, Chapter XIII, p. 685.

a series of sample rechecks showed the number of omissions to be negligible. A small number of cases of application for county hospital care were noted and withdrawn from the study, as it seemed inappropriate to include them in either the relief group or the control group. When a family applied for relief and also for county hospital care, the case was retained in the relief group; but the date used was that of application for relief; not for hospitalization.[1] Funds did not permit a detailed examination of each record to determine the date when the case became inactive, if at all. About two-thirds of the cases were inactive on July 1, 1934.[2] A "relief case" is defined, therefore, as a family to which public relief was granted regardless of the duration of the relief. Another definition would yield different results; the use of the present broad definition yields results which rather minimize than maximize the differences between the fertility of "relief" and "non-relief" families. All cases applying for relief before January 1, 1930, or after March 31, 1933, were discarded. About 100 cases were discarded because of lack of data on occupation or probable religion.

7. The admissible relief cases were next sorted into piles by year of marriage; by apparent religion ("Catholic" and "non-Catholic"), and by occupational group. The cards in the professional and proprietary group, being too few, were removed. This left $12 \times 2 \times 3 = 72$ sub-groups, with a total of 5,520 cards.

8. The non-relief cards (after subtracting the professional and managerial group and the cases of application for county hospital care but not for other public relief) were sorted into 72 corresponding piles. The control group was then obtained by drawing at random from each pile a group of cards, equalling the number in the corresponding relief file. This procedure served to control fairly well the factor of age as well as duration of marriage.

9. Each non-relief card was matched at random with a relief card in the corresponding sub-group, by entering on it a "beginning date" duplicating the date of application for relief on the relief card. The elapsed time in months from date of application of relief to December 31, 1933, was entered on each card.

[1] The "non-relief" cards probably include a very small number of cases which received relief from some private agency but never from the public agency. A random sample of 225 cards was checked at the Milwaukee Social Service Exchange. It was found that almost all cases which had been receiving private relief were transferred to public relief by the spring of 1933 and, therefore, appear in our "relief" group.

[2] A comparison of fertility of relief cases which were still active on July 1, 1934, with relief cases which became inactive shows that the former had a 60 per cent larger confinement rate than the latter. Such an excess probably is to be expected since the birth or expectation of a child would tend, on the average, to keep a family on the relief rolls, just as it would tend toward placing the family on the relief rolls for the first time. The interesting findings of Sydenstricker and Parrott, showing 66 per cent higher fertility among relief families, are perhaps more closely analogous to this comparison than to the main object of inquiry of the present paper. See *Milbank Memorial Fund Quarterly*. XII, April, 1934, especially Table 4.

10. The number of confinements occurring between 9 months after application for relief or the "beginning date" and December 31, 1933, was entered on each relief and non-relief card. Twins were counted as one confinement and stillbirths were considered for the purpose of the present study as equivalent to live births.

11. The months of exposure to risk of pregnancy were calculated uniformly for the two groups. It was assumed arbitrarily that a mother, on the average, would be free from further risk of pregnancy during a pregnancy period and during the first three months of lactation—total, twelve months. This period is perhaps too long, but a study of Table I will show that the choice of a shorter period would not materially alter the results. Subtraction further had to be made for removal of exposure to risk due to death or separation. The relief group, having more confinements, consequently had a shorter average exposure to the risk of pregnancy. In addition, pregnancy at the time of going on relief (or at the "beginning time" of the control group) shortened the period of exposure to risk, while the confinement in such an instance was not counted in our final figures. An example will illustrate the general procedure: Application for relief, March 25, 1931. Birth, October 1, 1931. Birth, July 28, 1933. Months elapsing between application and December 31, 1933: 33. Subtract for first confinement, $6+3=9$. Subtract for second confinement, $9+3=12$. Months of exposure to risk of pregnancy, 12. Number of confinements later than 9 months after going on relief, 1.

12. The sub-groups by year of marriage were consolidated and Table I was prepared.

The findings of the study may be summarized as follows:

1. The number of confinements in all sub-groups of occupational class and probable religion was greater among the relief families than the non-relief. Although the nature of the data does not facilitate an exact yet simple mathematical treatment, it is evident from a rough test, probably more conservative than an exact test, that the differences within the last three sub-groups in Table I are much too great to be attributed to chance. The significance of the other differences, taken individually, is doubtful, but the consistent *direction* of the excess of relief fertility in all six sub-groups argues quite indisputably for the significance of the difference in the *aggregate* number of confinements among relief and non-relief families in this study.[1]

[1] The crude test used is the calculation of $\chi = p_1 - p_2 / \sqrt{p_0 q_0 (1/n_1 + 1/n_2)}$ treated as a normal deviate. For example, in case of "Catholic" clerical workers, $\chi = \left(\frac{39}{206} - \frac{27}{206}\right) \Big/ \sqrt{\frac{66}{412} \times \frac{346}{412} \times \frac{2}{206}} = 1.7$, which is on the verge of significance, but not at all decisive. This formula would hold if each family

2. The excess of fertility in the relief group is greater if the calculations are based on months of exposure to the risk of pregnancy. If it is objected that the arbitrary subtraction of 12 months from the risk period is too great and tends to exaggerate slightly the differential between relief and non-relief families, it may be argued also that the true differential is even slightly greater than that shown in Table I, for the reason that no allowance has been made for reduction of risk period due to conceptions taking place after April 1, 1933, with confinements in 1934. To be meticulously logical, the present study should have been deferred until a complete report of births occurring up to October 1, 1934, was available. However, this would have necessitated corresponding again with all of the parents and deferring the report for perhaps nine months, while the gain in accuracy would almost certainly have been too slight to justify the delay. On the other hand, if the end period used in the study had been set back nine months, the large number of cases opened in the winter of 1932 could not have been used.

3. The fertility, in the brief period considered, of couples married by a Catholic priest, was higher in every sub-group than that of corresponding couples not married by a Catholic priest. These differences among the relief families may not be statistically significant; there hardly can be any question of the significance among the non-relief skilled, semi-skilled and unskilled workers. Incidentally, it is interesting to note that the "Catholic" fertility in every *non-relief* occupational group was lower than that of "non-Catholics" in the corresponding *relief* occupational group.

4. As usual in studies of differential fertility, the lowest fertility was in the clerical groups and the highest in the semi-skilled and unskilled worker groups.

This little study, if the generality of the results is confirmed by investigations made elsewhere, provides some factual basis for further interesting discussions of differential fertility, quite apart from its comparison of relief and non-relief families. In particular, data which reveal, even though crudely and unsatisfactorily, the degree of success of the efforts of the Catholic church to prevent the diffusion of the inventions of contraception among her adherents, will be of much interest to students of American population trends. A danger lies in the temptation to read too much into such data. The writer hopes that the present study's many limitations, which he has tried to set forth fully and accurately, will be kept in mind before hurrying to generalizations from it.

---

had an equal exposure to risk and if there was not more than one confinement per family. Neither assumption is true, but the systematic error probably is in the direction of minimising the significance of the true difference.

BOOLE'S CHALLENGE PROBLEM

By Edwin B. Wilson, *Harvard School of Public Health*

Boole's Challenge Problem is discussed by J. M. Keynes in his *Treatise on Probability*, pp. 186–189, where some of the history of the problem is given and a solution in terms of Keynes's special notation. It may be of interest to show that the problem yields very easily to the most elementary methods in customary notation. The problem is this:

The à priori probabilities of two causes A_1 and A_2 are c_1 and c_2 respectively. The probability that if the cause A_1 occur, an event E will accompany it (whether as a consequence of A_1 or not), is p_1, and the probability that E will accompany A_2, if A_2 present itself, is p_2. Moreover, the event E cannot appear in the absence of both the causes A_1 and A_2. Required the probability of the event E.

A formulation in terms of the concept of association and the four-fold table (Yule, *Introduction to the Theory of Statistics*, Chap. III) appears helpful. There are two tables, one for the causes and one for the event E.

	A_1 present	A_1 absent	Totals	Occurrences of E		
A_2 present.....	α	γ	$\alpha+\gamma$	α_1	γ_1	$\alpha_1+\gamma_1$
A_2 absent.....	β	δ	$\beta+\delta$	β_1	0	
Totals.....	$\alpha+\beta$	$\gamma+\delta$	"N"	$\alpha_1+\beta_1$		

If we resolve the probabilities into frequencies, we see in the first table that out of N trials both A_1 and A_2 occur α times and neither occurs δ times, etc. Hence the given probabilities are

$$c_1=(\alpha+\beta)/N, \qquad c_2=(\alpha+\gamma)\ N \tag{1}$$

It is well known that to determine a four-fold table (association) four independent numbers, or four probabilities whose sum is unity, are requisite. As only two probabilities c_1 and c_2 are given, there must be one degree of freedom in the probabilities of the association table for the causes A_1, A_2. The second table enumerates the occurrences of E for each place in the table for A_1, A_2. Obviously

$$0\leqq\alpha_1\leqq\alpha, \quad 0\leqq\beta_1=\beta, \quad 0\leqq\gamma_1=\gamma \tag{2}$$

$$p_1=(\alpha_1+\beta_1)/(\alpha+\beta), \quad p_2=(\alpha_1+\gamma_1)/(\alpha+\gamma) \tag{3}$$

As only two probabilities p_1 and p_2 are given in the statement of the problem whereas three independent numbers α_1, β_1, γ_1, may occur, it is

clear that the data concerning E have also one degree of freedom, that the problem therefore possesses two degrees of freedom, and that its solution should involve two variables which are freely assignable, at least within certain limits.

Now the probability of E is, by definition,

$$P=(\alpha_1+\beta_1+\gamma_1)/N=c_1p_1+c_2p_2-\alpha_1/N \tag{4}$$

and depends apparently on only one unknown, viz., the ratio α_1/N. Let it be assumed that P is known.

We have

$$\alpha_1/N=c_1p_1+c_2p_2-P,\ \beta_1/N=P-c_2p_2,\ \gamma_1/N=P-c_1p_1$$

and hence conditions (2) impose on P the conditions

$$c_1p_1+c_2p_2-P\geqq 0,\quad P-c_2p_2\geqq 0,\quad P-c_1p_1\geqq 0$$

Further, as identically

$$\alpha/N=1-\beta/N-\gamma/N-\delta/N=c_1+c_2+\delta/N-1$$

$$\beta/N=1-c_2-\delta/N,\quad \gamma/N=1-c_1-\delta/N$$

conditions (2) also yield

$$P-c_1p_1\leqq 1-c_1-\delta/N,\qquad P-c_2p_2\leqq 1-c_2-\delta/N$$
$$c_1p_1+c_2p_2-P\leqq c_1+c_2+\delta/N-1$$

From this it follows that P must satisfy the relations

$$\left.\begin{matrix} c_1p_1 \\ c_2p_2 \\ c_1p_1+c_2p_2-c_1-c_2-\delta/N+1 \end{matrix}\right\}\leqq P\leqq\left\{\begin{matrix} c_1p_1+c_2p_2 \\ c_1p_1+1-c_1-\delta/N \\ c_2p_2+1-c_2-\delta/N \end{matrix}\right. \tag{5}$$

The inequalities (5) imply nine inequalities because each of the three expressions on the left must be less than or equal to each of the three on the right, of which some are identically satisfied—thus $c_1p_1\leqq c_1p_1+c_2p_2$. Moreover δ/N cannot be negative. Hence limits for δ/N are as follows

$$0,\ 1-c_1-c_2\leqq\delta/N\leqq\left\{\begin{matrix} 1-c_1,\ 1-c_2+c_2p_2-c_1p_1, \\ 1-c_2,\ 1-c_1+c_1p_1-c_2p_2 \end{matrix}\right. \tag{6}$$

Of the eight inequalities arising from the fact that either of the two values on the left must be less than any of the four on the right only two place restriction on c_1, c_2, p_1, p_2, viz.,

$$c_1p_1-1+c_2\leqq c_2p_2\leqq c_1p_1+1-c_1 \tag{7}$$

or

$$c_2p_2-1+c_1\leqq c_1p_1\leqq c_2p_2+1-c_2 \tag{7'}$$

If then we have values of c_1, c_2, p_1, p_2 which satisfy (7), or its equivalent (7′), and values actually arising must satisfy it, we may find from (6) a value, and in general a range of values, for δ/N and from (5) a value, and in general a range of values, for P, the two-fold in determinacy of which is thus exhibited; and from these values of δ/N and P we may then find α/N, β/N, γ/N, α_1/N, β_1/N, γ_1/N so that in terms of δ/N and P the whole set-up of the association tables is fixed so far as concerns the probabilities and becomes fixed in respect to the numbers if the total population N of trials be known.

From (6) we have

$$c_1p_1+c_2p_2-c_1c_2+1-\delta/N \leqq \begin{cases} c_1p_1-c_2(1-p_2),\ 2c_1p_1-c_1 \\ c_2p_2-c_1(1-p_1),\ 2c_2p_2-c_2 \end{cases}$$

Hence if δ/N be taken as large as the least of the four quantities on the right of (6) the expression on the left of the above inequality is equal to the least of the four quantities on the right, and as each of them is less than c_1p_1 or less than c_2p_2, the quantity on the left must be less than c_1p_1 or c_2p_2; thus in the absence of a specified value for δ/N, one may say that there are possible values of δ/N for which the third item on the left of (5) may be suppressed, as it is in the table given by Keynes (*l.c.*; p. 189), but there may be values of δ/N small enough so that this expression is greater than c_1p_1 or c_2p_2. If we make a plot of P against δ/N the conditions (5) show that, for a given δ/N, P is limited to a portion of a (vertical) line and that as δ/N varies the extremities of the line move along horizontal lines or lines inclined at 135° to the axis of δ/N; thus the locus of possible pairs of values of P and δ/N is a figure bounded by horizontal or vertical lines or lines inclined at 135°.

As one example, suppose $c_1=\frac{1}{3}$, $c_2=\frac{2}{3}$, $p_1=\frac{1}{2}$, $p_2=\frac{1}{3}$ Conditions (7) are satisfied, (6) gives $0\leqq\delta/N\leqq\frac{1}{3}$, and (5) gives

$$\left(\frac{1}{6}\right),\ \frac{2}{9},\ \frac{7}{18}-\frac{\delta}{N}\leqq P\leqq\frac{7}{18},\left(\frac{5}{6}-\frac{\delta}{N}\right),\ \frac{5}{9}-\frac{\delta}{N}$$

Clearly the two terms in parentheses are ineffective and may be discarded. If $\delta/N=0$ there is a single solution $P=7/18$; if $0<\delta/N\leqq 1/6$, P has the range $7/18-\delta/N\leqq P\leqq 7/18$; if $1/6\leqq\delta/N<1/3$, P has the range $2/9\leqq P\leqq 5/9-\delta/N$; and if $\delta/N=1/3$ there is a single value of $P=2/9$. The points $(\delta/N, P)$ fill a parallelogram bounded by the horizontal lines $P=7/18$, $P=2/9$ and by the inclined lines $P=7/18-\delta/N$ and $P=5/9-\delta/N$.

Other examples may be worked out from the assumptions

$$c_1=\tfrac{2}{3},\ c_2=\tfrac{3}{4},\ p_1=\tfrac{1}{2},\ p_2=\tfrac{1}{3}$$
$$c_1=\tfrac{1}{2},\ c_2=\tfrac{2}{3},\ p_1=\tfrac{2}{3},\ p_2=\tfrac{3}{4}$$
$$c_1=\tfrac{1}{2},\ c_2=\tfrac{2}{3},\ p_1=\tfrac{1}{2},\ p_2=\tfrac{1}{3}$$

to find a variety of shapes of regions for the locus of $(\delta/N, P)$.

The solution has been given by including P within inequalities which contain a parameter. If it were preferred to give the value of P by an equation, (4) could be used and then α_1/N could be included between inequalities obtainable directly from (5).

One may well wonder what importance there is, other than that of satisfying intellectual curiosity, in stating or solving an indeterminate problem. If the problem arises in practice the statistician should realize that he needs to accumulate counts not only of $\alpha+\beta$ and $\alpha+\gamma$ but of α, not only of $\alpha_1+\beta_1$ and $\alpha_1+\gamma_1$ but of α_1. It is indeed difficult to see how a person should be able to say that on N trials E accompanies $A_1, \alpha_1+\beta_1$ times and $A_2, \alpha_1+\gamma_1$ times and is absent if A_1 and A_2 are both absent without being able also to make the enumeration of the α_1 times when E occurs with both A_1 and A_2; and so, too, if we know the probability of A_1 and of A_2 individually we should ordinarily, except for carelessness, be able to give the probability of both A_1 and A_2 together —whether à priori or à posteriori. Any statistical consultant knows, however, from his own experience that investigators not trained in the logic of statistics do often come to him with only partial enumerations of the necessary elements to make their problem definite and cannot always go back and pick up the missing information; it then becomes necessary to do the best one can with the data available.

INDEX OF PRODUCTION OF MANUFACTURES DERIVED FROM CENSUS DATA, 1929, 1931

By Aryness Joy

The comprehensive index of production of manufactures compiled from census data by E. E. Day and Woodlief Thomas, and published in the census monograph, *The Growth of Manufactures, 1899–1923*, has recently been computed for the years 1929 and 1931. Figures for the year 1927 were published in this Journal for December, 1930, together with a description of the index. The biennial index numbers on a 1919 base for the period 1914 to 1931 are as follows:—

Index of Production of Manufactures

(1919=100)

1914	79.3	1925	126.9*
1919	100.0	1927	127.7*
1921	79.2*	1929	144.8
1923	121.5*	1931	96.0

* Revised.

This index includes 130 series representing directly or indirectly the physical volume of output in 53 industries and 13 industrial groups, which make up the major portion of manufacturing production. The quantity data, which for the most part represent semi-finished goods, are derived almost entirely from the biennial reports of the Census of Manufactures, with the exception of certain statistics collected by the Bureau of Internal Revenue, the Bureau of Mines and other governmental agencies. Because of the comprehensive nature of census reports, an index computed from these data affords the most reliable indication of changes in the physical volume of manufacturing output. It is especially valuable for certain industrial groups, such as foods and chemicals, in which current reports of production are not available for important branches of the industry.

In general, this index confirms the indication of changes in output given by monthly indexes of production, for example, that compiled by the Federal Reserve Board. From 1927 to 1929 the Census index indicated an increase of 13.3 per cent and the Federal Reserve Board's index an increase of 12.1 per cent. From 1929 to 1931 the two indexes showed decreases of 33.7 per cent and 31.8 per cent, respectively. For certain industrial groups, however, differences are considerably larger, the variations offsetting each other in the total index. These group indexes are shown in the table at the end of this article.

The same formula used by Day and Thomas in the original index has been followed, with the exception of the index for automobiles, in which modifications were made.[1] The index is a weighted geometric average of relatives on a 1919 base. The weights represent an average of the relative importance of the industry or industrial group in the base year and the given year. Weights for industrial groups and particular industries were derived from census figures measuring the value added by the process of manufacture, and for individual products within each industry from statistics of total value of products. Weights were assigned to each of the 13 major industrial groups in proportion to their importance in the total value added by the process of manufacture for the groups represented, with modifications determined by the extent to which available data were representative of output of other products not directly represented in the index by volume figures. Thus, steel production was given a weight representing value added by manufacture, not only for products of "steel works and rolling mills" but also for all other industries in the iron and steel group with the exception of blast furnaces, which were represented by pig iron, and one-half the total for automobiles and ship-building. These group weights were distributed among component industries in proportion to the importance of each in the group of industries included in the index, and the weight as assigned to each industry was divided among its individual products in similar fashion.

The use of a group weighting system represents the most important difference between the index compiled by Day and Thomas and that compiled from census data by Frederick Mills and published in *Economic Tendencies in the United States*. Because of the varying proportion of output represented by quantity figures in different industrial groups, the use of a simple set of direct weights gives very different results from the use of a group weighting system. A second index of production based on census data and compiled by the National Bureau of Economic Research also employs a group weighting system.

In computing the index for 1931 the weights varied more widely than those used in earlier census years, except in 1921, as a result of the declining importance of heavy manufacturing industries, compared with consumers' goods industries.

Further details regarding series included and weights used will be supplied on request.

[1] Revisions in the index for motor vehicles consist of substituting a simple count of number of automobiles for an index which assigned varying weights to open and closed passenger cars and other vehicles. The revised figures for the years 1921, 1923, 1925 are as follows:

	1921	*1923*	*1925*
Motor vehicles	84.4	206.2	220.8
Vehicles for land transportation, group	71.9	181.0	182.3

INDEX OF PRODUCTION OF MANUFACTURES
(1919 = 100)

Industry	1927	1929	1931
ALL MANUFACTURING	127.7	144.8	96.0
I. FOOD AND KINDRED PRODUCTS	118.1	128.8	118.7
Butter, cheese, and condensed and evaporated milk	145.0	150.5	149.9
Slaughtering and meat packing	108.0	113.9	108.3
Flour-mill and grain-mill products	88.1	87.0	81.1
Sugar, beet	121.7	136.1	147.6
Sugar refining, cane	145.6	145.3	125.8
Canning and preserving, fruits and vegetables, etc.	143.1	176.6	153.2
Canning and preserving, fish, oysters, etc.	106.8	146.6	82.5
Corn syrup, corn oil, and starch	131.9	142.9	101.1
Rice cleaning and polishing	117.3	117.0	118.8
Ice, manufactured	144.2	162.7	154.8
II. TEXTILES AND THEIR PRODUCTS	125.7	130.7	107.5
Cotton goods	133.2	127.9	103.3
Knit goods	115.2	130.4	113.1
Silk manufactures	143.2	173.8	159.4
Woolen and worsted goods	106.2	107.6	89.3
Carpets and rugs, wool, other than rag	128.8	140.7	84.7
Cordage and twine	94.8	102.4	70.6
Hats, fur-felt	97.7	84.3	62.2
Oil cloth and floor coverings	242.9	265.7	182.6
III. IRON AND STEEL AND THEIR PRODUCTS	127.0	157.5	70.9
Blast furnaces	118.6	139.1	59.9
Steel works and rolling mills	127.8	159.2	71.9
IV. LUMBER AND ITS REMANUFACTURES	102.4	108.4	48.0
Lumber and timber products	102.4	108.4	48.0
V. LEATHER AND ITS FINISHED PRODUCTS	98.3	103.0	86.4
Leather, tanned, curried, and finished	86.7	84.4	68.2
Boots and shoes	103.7	114.2	97.3
VI. PAPER AND PRINTING	165.7	181.3	152.7
Paper and wood pulp	155.3	169.0	142.2
Printing and publishing	167.2	181.1	151.9
Paper boxes	202.1	238.4	206.1
VII. CHEMICALS AND ALLIED PRODUCTS	157.8	186.7	146.8
Chemicals and acids	155.6	204.6	163.0
Coke	115.2	135.0	75.5
Gas	129.5	128.5	117.4
Fertilizers	97.5	112.1	84.0
Explosives	108.5	104.7	69.3
Paint and varnish	168.2	188.0	130.3
Petroleum refining	240.1	316.4	275.7
Oil and cake, cottonseed	124.9	108.6	98.0
Turpentine and rosin	185.4	182.9	140.9
Salt	110.0	124.1	106.9
Soap	120.9	122.6	117.2
VIII. STONE, CLAY, AND GLASS PRODUCTS	175.6	174.7	104.7
Brick and tile, terra cotta, and fire-clay products	156.8	138.0	61.0
Lime	134.3	130.6	95.5
Cement	215.6	212.7	155.8
Glass	171.2	195.5	123.6
IX. METALS AND METAL PRODUCTS, OTHER THAN IRON AND STEEL	129.7	158.5	90.5
Brass, bronze, and other non-ferrous alloys	124.2	196.4	99.7
Smelting and refining	126.6	131.2	74.9
Other industries	134.1	149.1	92.2
X. TOBACCO MANUFACTURES	148.5	155.0	149.0
Tobacco, chewing and smoking, and snuff	93.4	89.9	87.5
Tobacco, cigars	92.2	92.2	75.6
Tobacco, cigarettes	187.9	230.4	220.4
XI. VEHICLES FOR LAND TRANSPORTATION	145.8	210.3	83.9
Motor vehicles	177.4	281.0	122.0
Carriages, wagons, and repairs	17.8	16.7	4.4
Railway equipment	56.6	54.6	10.6
XII. SHIP AND BOAT BUILDING	8.8	8.2	7.9
XIII. RUBBER PRODUCTS	171.4	180.4	121.8
Rubber boots and shoes	118.0	107.8	75.1
Rubber tires and inner tubes	184.1	195.7	131.0

ON THE RAPIDITY OF THE DECLINE IN THE BIRTH RATE

By P. K. Whelpton, *Scripps Foundation for Research in Population Problems*

Most persons who give any thought to population matters in the United States realize that the birth rate has been declining. It appears, however, that the great rapidity of the decline has not yet been appreciated even by some careful students in this field.

In his excellent study *Introduction to the Vital Statistics of the United States, 1900 to 1930*, Willcox writes:

> There is a strong presumption that more than one-fifth of the children under 5 at the time of any census were under 1. That was true in 21 out of 22 European countries, all but Greece, for which post-war census figures are at hand, and in 16 out of 21 European countries, all but Ireland, Spain, Portugal, Switzerland, and Bulgaria, for which pre-war censuses are accessible. In view of the uniformity of birth rates through a short series of years the probability that the United States is an exception can almost be disregarded. [1]

As a matter of fact, however, not only the birth rate but the number of births in the United States has been declining so rapidly that this nation in all probability *was* an exception in 1930.

The birth registration area of 1925 included 33 states and the District of Columbia. In this area there were 9,084,561 births registered during 1925 to 1929, and 1,729,074 during 1929. Births during 1929 thus constituted 19.03 per cent of the five-year total in this area. Studies made by the writer indicate that when allowance is made for births in nonregistration states and unregistered births in the registration area, there were 13,467,000 births during 1925–29, and 2,526,000 during 1929, the latter being 18.75 per cent of the five-year figure.[2] These percentages are in close agreement, the difference of .28 per cent being attributable in part to interstate migration, and to a more rapid decline in the birth rate in some states than others. It appears certain, therefore, that births in the United States during 1925–29 were more than 5 times as numerous as during 1929, as was indisputably the case with registered births in the 1925 birth registration area.

The probable number of children under 5 on January 1, 1930, surviving from births during 1925–29 may be calculated by means of life table values. The writer has used death rates from life tables of the

[1] Walter F. Willcox, *Introduction to the Vital Statistics of the United States, 1900 to 1930*. Washington. G. P. O., 1933 (Bureau of the Census), p. 81.

[2] These figures are slightly different from those published in P. K. Whelpton, "Population," *The American Journal of Sociology*, Vol. XXXVIII, No. 6, May, 1933, p. 826, because of more recent data.

Metropolitan Life Insurance Company,[1] converting them from a population in which all births occur at the beginning of each year to one in which births are distributed evenly throughout the year by means of the percentages used by Miss Foudray in *United States Abridged Life Tables, 1919–1920*, p. 36. Applying these death rates to the number of births in the 1925 birth registration area indicates that there were surviving on January 1, 1930, 8,511,000 children under 5 and 1,661,000 children under 1, the latter being 19.5+ per cent of the former. Using the above estimates of births for the United States, these death rates indicate 12,617,000 children under 5 and 2,426,000 or 19.2+ under 1.

So much for the situation on January 1, 1930. Although the trend of births from 1924 to 1933 was decidedly downward, births were more numerous during the first 3 months of 1930 than in the corresponding period of 1929. Reflecting this situation, the writer's estimates of births during five- and one-year periods ending on the census date April 1, 1930, are 13,395,000 and 2,546,000, respectively. Applying the same death rates as before gives 12,549,000 children under 5 on the census date, with 2,446,000 or 19.5− per cent under 1. Children under 1 would thus be expected to constitute less than one-fifth of the children under 5 in the native population in the last census even if the enumeration of children under 1 was as complete as that of children under 5, which is not usually the case. It is admitted that the number of births and the death rates used herein for the United States are not absolutely exact—such data do not exist for the entire nation—but the adjustments for unregistered births in registration states and for births in states not in the registration area have been made on a uniform basis from year to year, and the death rates taken from unbiased authorities, hence the percentage 19.6− should be very close to the truth.

And the rapid decline of births in the United States is yet unchecked!

[1] Metropolitan Life Insurance Company, *Statistical Bulletin*, Vol. X, No. 2 (February, 1929), p. 2, and photostat pages received from A. J. Lotka.

LONDON MEETING OF THE INTERNATIONAL STATISTICAL INSTITUTE

By Walter F. Willcox

The Royal Statistical Society, founded in 1834 as the Statistical Society of London, celebrated its centenary in April of this year. At its jubilee meeting, postponed from 1884 to 1885, it had arranged for a statistical congress to discuss the revival of international statistical meetings which had begun in the Brussels congress of 1853 and, after eight more sessions, had come to an end as a result of international jealousies entailed by the Franco-Prussian War. The discussion in 1885 resulted in the establishment of a private or unofficial organization, the International Statistical Institute, with membership narrowly limited but drawn mainly from the group of officials. The Royal Statistical Society, the parent of the Institute, naturally invited its off-spring to the centenary, an invitation which was gladly accepted.

The sessions began Monday morning, April 16, and continued until Friday noon. The business meetings, starting after the formal opening session, were divided, as usual, into three sections under the chairmanship of Professors Gini and Rappard and of Sir William Beveridge, the first section considering demographic and mathematical papers, the second economic papers, the third papers on more general social questions. The proceedings will soon be published in full, probably before the end of the present year. In view of that prospect it is needless to summarize the papers in this note.

There were 45 communications, of which two were from nonmembers.

The present number of members, the attendance and the number of communications from the leading countries were:

Country	Number of members	Attendance	Communications
France	24	10	5
Great Britain	23	13	4
Germany	21	10	5
United States	20	2	7
Italy	17	8	10
Hungary	7	4	4
Poland	7	1	1
Belgium	6	2	1
Holland	6	4	2

Of the American members only E. D. Durand and the writer attended, but the American contributions, with papers by Messrs.

Fisher, Mitchell, Pribram,[1] Snyder (two communications), Sydenstricker and Kuznets, were worthy of the notable occasion.

Owing to the founding a century ago of two statistical societies within a year of each other, the present centenary closely followed that of the Mexican Society of Geography and Statistics. The generous hospitality of these societies to the International Institute brought its last two sessions close together instead of at the usual two-year interval.

At the Mexico meeting the most important discussion had been over Mr. Fisher's paper on "The Debt Deflation Theory of Great Depressions." This ended in the appointment of a committee or "commission," of which Mr. Pribram became chairman and Messrs. Mitchell, Fisher and Snyder American members.[2] The chairman had secured from the United States four communications in addition to his own. As a result American participation at London was comparable with that of members from the leading European countries; at the sessions of 1933 and 1934 together it exceeded that at the thirteen sessions between 1885 and 1913 (omitting the one at Chicago).

The attendance and participation of American members are less common than of European. This is due mainly to the greater cost of the journey, to the inability of the American Government without specific Congressional action to pay expenses even when members are duly accredited by the Department of State, and to the prevalent idea that membership in the Institute is primarily an honor rather than an opportunity. The following figures show for the leading countries the per cent that the average number attending one of the eleven meetings since 1900 (excluding three at which members' expenses were partly or wholly borne by the host) makes of the present number of members:

France	55	Germany	41	Switzerland	25
Hungary	50	Netherlands	40	England	19
Belgium	50	Denmark	35	United States	8
Sweden	45	Italy	25		

They show that American members have participated less commonly than those from any of the European countries for which the facts have been obtained.

The question of revising its statutes has been before the Institute almost ever since 1923, when its meetings were renewed. At London it was satisfactorily and harmoniously closed; a few minor changes were adopted, perhaps the most important being an increase in the number

[1] As Mr. Pribram now lives in the United States and his paper was based on his recent studies, it has been classed with the American contributions.

[2] For the resolution referred to this committee, see this JOURNAL for March, 1934, p. 68.

of members admissible from 220 to 250 and a reduction in the number from any one country from one-fifth of the total to 25.

The centenary of the Royal Statistical Society was observed at a meeting addressed by its President, Lord Meston, its Honorary President, the Prince of Wales, and the President of the Institute, Professor Zahn; and later at a banquet followed by speeches from the President of the Royal Society, the cradle of modern statistics, a former President of the Royal Statistical Society, the President of the Institute, and others.

Delightful social functions included a reception at the London School of Hygiene and Tropical Medicine, and receptions by the Lord Mayor and Sheriffs of London at the Guildhall, and by the Government at Lancaster House. The guests were entertained also by a gala performance at the theater.

On Saturday morning a large group of members journeyed to Dublin, where they were welcomed by President de Valera, entertained at a banquet by the Statistical and Social Inquiry Society of Ireland, given a glimpse of the beauties of the Emerald Isle, and charmed by the hospitality of its citizens.

TAMING THE WEEKLY SEASONAL

By G. M. Whitright and M. Farrington

Weekly figures for the various phases of business and financial activity, even when seasonally corrected, have a most disconcerting habit of bouncing in one direction for one week and of pitching in the opposite direction in the week immediately following. The problem of extracting from these erratic movements the maximum amount of reliable information is an important one for the economic statistician during these days of rapid changes in business activity.

It takes time to develop a trend, and a single week is as likely to mislead one as it is to give any reliable indication of even a short time trend in business swings which currently are of interest to leaders of industry and finance. Improvement in the business of a corporation for a period of six months is of fundamental importance because betterment over so long a period is reasonable proof that the trend is upward. Moreover the fact that the accounting year is half over makes it unlikely that any easing of activity during the next six months will have any appreciable dampening effect on the earnings for the full year. Some significance may also be attached to favorable reports for a quarter of a year, since improvement of this duration indicates that gains for one month have been substantiated by further gains of sufficient magnitude to make the quarterly reports favorable.

Thus we see that even monthly figures have, at times, little significance to students of business unless they are carefully compared and combined with the monthly figures of the previous period of several months. But weekly figures are being published in increasing numbers, and it is our practical problem to give them their proper significance and to find a useful place for them in our analysis of economic conditions. One of the best ways of evaluating weekly figures is to consider them as advance indications of the results of the complete month, which, of course, will be then analyzed in terms of the previous several months.

This article is written from the point of view of the student who has had some experience in seasonal corrections of monthly figures of business activity, such as electric power production and car loadings. In brief, if the monthly figure for electric power production for December of each year is normally 4 per cent greater than the average month of the year, the total electric power production figures for the current month of December is reduced by 4 per cent in order to determine

whether or not the activity of this month deprived of its usual seasonal activity was higher or lower than the similarly corrected figure for November. It is upon these seasonally corrected monthly figures that the seasonal indexes for weekly figures are based.

To discuss the various methods of analyzing weekly figures that are now being currently used would require the publication of a fair sized book that would involve the reader in a maze of calculations, assumptions, complicated mathematical refinements, and calendar irregularities that are not warranted here. If the reader has had any experience with correcting weekly figures for seasonal variation, he is thoroughly familiar with the difficulties which are inherent in the problem. It might be noted in this connection that many of the analysts' worries over weekly and monthly figures might be considerably lessened by the introduction of a new national calendar in which the weeks, months, and holidays occurred with greater regularity and with more consideration for the student of business and his problems.

Recalling that the primary purpose of analyzing weekly figures is to find some indication of what the results of the month are going to be before the monthly figures become available, we find that the purpose itself suggests a simple method of dealing with the weekly figures. In essence the method is this:

1. Select a series of monthly figures for the year previous to the current year upon which you are working; obtain the monthly seasonal indexes for this series; and apply these indexes to the monthly totals, arriving finally at a seasonally corrected figure for each month of the year.

2. Arrange the corresponding weekly series in chronological order and calculate a four weeks' moving total. In other words, opposite each of the actual weekly figures you will have a figure computed by adding to that week the sum of the three previous weeks. You have here what approximates a moving month, containing just four weeks, for each of the fifty-two weeks of the year.

3. Take the four weeks' moving total figure corresponding to the last week in each month and divide it by such a figure (found by simple algebra) that you will arrive at the identical result obtained by correcting the original monthly total corrected for seasonal variation. This new figure which you use as the divisor for the four weeks' moving total of actual figures may be termed the new seasonal index for what could be called a "sliding month."

4. The calculation made above will result in one seasonal correction index for each month of the year, with the index to apply to the last four weeks' moving total in each month. The fourth step is to obtain

seasonal indexes for the weeks falling between these twelve points. This is done by what is called straight-line interpolation, and this process is adequately described in practically any good elementary text on statistics.

5. You now have 52 seasonal indexes for your test year, say 1932, and these can be applied directly (by division) to the current 1933 series of weekly figures which you have put in the form of four weeks' moving totals.

One of the advantages of this method is its relative simplicity as compared with most methods now in use. The results fulfill the desired purpose of indicating, without false moves, the direction of the monthly trend so that by the third week of the month at least a fairly reliable estimate of the complete month can be obtained.

Furthermore, the use of the four weeks' moving total, which in itself is a smoothing appliance, makes it unnecessary to worry about such factors as whether the Fourth of July falls in the 26th week or in the 25th week of the year, as compared with what was the case last year. The determination of the seasonal index in such a way that the result once a month is in agreement with the monthly seasonally corrected figure forms a convenient manner by which to tie together two series which are not strictly comparable when they appear as monthly figures and when they appear as weekly figures. For example, the monthly electric power series, used in testing this method, was expressed in terms of a base period and was originally published on the average daily basis, while the weekly figures were expressed in terms of actual totals and included some small items not contained in the monthly figures. For instance, the monthly figures did not include imports from Canada, but this small difference (which may be assumed to be constant over relatively short periods of time) was taken care of when one weekly index for each month was so calculated that it gave the same result as the monthly seasonally corrected figure.

Weekly figures are little more than news items, and they are significant to the economic statistician only in so far as they give an indication of what the final results of the month will be. The method given here may be classified as one in which mathematical and theoretical rigor are given secondary consideration in order to simplify the problem to limits justified by the original purpose and the character of the data to be analyzed. This method is of value mainly in series of weekly data whose four weeks of moving totals exhibit relatively small amplitude of fluctuation about the straight lines connecting the end-of-the-month points. It is primarily concerned with the trend of developments through the weeks.

ON A CERTAIN CONCLUSION OF FRISCH'S

By S. Subramanian, *Annamalai University, South India*

1. Writing on the "Necessary and Sufficient Conditions Regarding the Form of an Index Number Which Shall Meet Certain of Fisher's Tests," [1] Frisch works out a certain form for the index number which satisfies the base test and the commensurability test at the same time and from this form he concludes that the base test, the commensurability test, and the determinateness test are incompatible (the index number having continuity and differentiability properties). The purpose of this paper is to point out how the above conclusion is not quite correct. Almost throughout, I follow Frisch's notation.

2. If p_t is zero, considering that particular commodity, we have

$$P_{ts}=f(p_s;\ x_t,\ x_s).$$

The commensurability test easily leads to

$$P_{ts}=\phi(x_s/x_t,\ a_s) \tag{1}$$

whereas when $p_t \neq 0$,

$$P_{ts}=\Phi(p_t/p_s,\ a_s,\ a_t) \tag{2}$$

So the form of P_{ts} has been entirely changed.

3. If P_{ts} is to satisfy the base test also, we have

$$\text{if } \psi=\log \phi$$

$$\frac{\partial^2\psi}{\partial_s\partial_t}=0.$$

This on simplification gives

$$\psi_1+\psi_{11}\, x_s/x_t=0,\ 6,\ \psi=A(p_sx_s)+c\log x_s/x_t.$$

We get now

$$P_{ts}=B(p_sx_s)\ .\ (x_s/x_t)^c \tag{3}$$

When p_t does not vanish, Frisch has proved that

$$P_{ts}=U(p_tx_t)+V(p_sx_s)+c\log p_t/p_s \tag{4}$$

Moreover, his form (19) has been derived on the basis of form (4) above, that is, on the assumption that no p vanishes. Hence no p can be made to tend to zero in his form (19). If any p vanishes, then, changes

[1] This Journal, Vol. 25 (1930), p. 404.

must be made in his form (19) in accordance with form (3) above. Similar changes will take place if any p_s, x_t, or x_s becomes zero.

4. In view of what has been said in the previous section, it is not correct to conclude that the base test, the commensurability test and the determinateness cannot be fulfilled at the same time. In fact, what can be said is only that Frisch's formula (19) is all right so long as no p or x vanishes. But when some p or x vanishes his formula (19) is not quite true but has to be modified according to the law given in the previous section.

REPLY TO MR. SUBRAMANIAN'S NOTE

By Ragnar Frisch, *University of Norway*

The point raised by Mr. Subramanian will be immediately cleared up when we state precisely what is meant by the condition $p_t \neq 0$ (or any of the other parameters p_s, x_t or $x_s \neq 0$).

In index number work we meet, of course, occasionally situations where some particular price or quantity is zero. It is therefore wanted to use formulae that do not become indeterminate in these cases, this is just the content of the determinateness test. My argument in the paper quoted by Mr. Subramanian is *not* dependent on the assumption that such situations never occur in the data to which it is proposed to apply the formulae studied. If this should have been the case, it would have meant a rather severe limitation of the applicability of the formulae obtained.

On the other hand, if—for some special reason—we should decide to work with an index number formula where some special price or quantity is *always* put equal to zero, in other words if we decide to work with a formula that does not at all depend on this price or quantity, then—and only then—will Mr. Subramanian's argument apply. It seems to me, however, that this is a rather trivial limitation of the validity of the formulae because in this case one would simply reduce the number of variables correspondingly, and to the new set of variables (none of which is now subject to the condition of always being equal to zero) my formulae would apply unconditionally.

A STATISTICAL SLIDE RULE

At last the statistician, particularly in the fields of education and psychology, may be on a parity with the engineer when it comes to facilitating computation. A study [1] which has been under way for two years has resulted in the production of a genuine calculating slide rule adapted to the statistical calculations of education and psychology. This statistical slide rule, models of which have been produced by Keuffel and Esser Company, is comparable to the engineer's log log duplex in versatility and number of scales. Ten of the seventeen scales are standard slide-rule scales (A, B, C, D, CF, DF, CI, CIF, K, L), which take care of the usual processes of multiplication, division, powers, roots, reciprocals, and logarithms. The remaining seven scales make the statistical slide rule virtually "self-contained" for ordinary statistical computations. They enable the user to read either $1-r^2$ or $\sqrt{1-r^2}$ for a given setting of r, to read areas under the normal curve for given abscissae, or *vice versa*, and to calculate intelligence or other quotients based on two ages given in years and months. The gauge point (.)6745 is marked on the C and D scales, including their variants.

An abstract of the study, with cuts of both faces of the statistical slide rule, appears in the July, 1934, issue of the *Peabody Journal of Education*. This abstract gives a description of the scales and summarizes the research which was made in attempting a tentative evaluation of the rule. It is predicted that the statistical slide rule will provide a valuable *vade mecum* for the individual student or research worker, including, of course, the educational psychologist.

CORRECTIONS

I

The Editor, Journal of the American Statistical Association

Dear Sir:

My attention has been called to two balancing errors in the 1910 numbers for males in Table 2 of my article on "A Social-Economic Grouping of the Gainful Workers of the United States," in the December, 1933, JOURNAL. These two errors involve other errors in the numbers and the percentages for the total and for males.

In the 1910 section of Table 2, the figures for Groups 1, 2, 2-c, and 4 should be as follows:

Group	Total		Male		Female	
	Number	Per cent	Number	Per cent	Number	Per cent
1. Professional persons	1,632,638	4.3	913,866	3.0	718,772	8.9
2. Proprietors, managers, and officials	8,579,746	22.5	8,183,563	27.2	396,183	4.9
2-c. Other proprietors, managers, and officials	1,201,289	3.1	1,146,276	3.8	55,013	0.7
4. Skilled workers and foremen	4,364,060	11.4	4,267,327	14.2	96,733	1.2

[1] E. R. Enlow, *A Statistical Slide Rule*. Contributions to Education, No. 130, George Peabody College for Teachers, Nashville, Tennessee, 1934. Unpublished doctoral dissertation. Abstract in the *Peabody Journal of Education*, July, 1934.

I think that the numbers for 1910 probably will be used but little and that the principal use that will be made of the percentages for 1910 will be as a measure of the change that took place from 1910 to 1920, or from 1910 to 1930. While the corrections in the published figures involve large numerical changes, the largest change in any percentage is 0.3.

Of course, I regret these errors very much. I hope they will not detract measurably from the usefulness of the grouping presented in my article.

Sincerely yours,

ALBA M. EDWARDS

II

The Editor, Journal of the American Statistical Association

Dear Sir:

In my note on "A Coefficient of Independent Determination" that appeared in this JOURNAL, March, 1934, the regression coefficient was omitted from the formula for $\bar{x}_1$. The correction follows:

$$\bar{x}_1 = b_{1(2-3)}\ (x_2 - b_{23\cdot4}x_3 - b_{24\cdot3}x_4)$$

$$S\bar{x}^2_1 = b_{1(2-3)}\ (Sx_1x_2 - b_{23\cdot4}\ Sx_1x_3 - b_{24\cdot3}\ Sx_1x_4)$$

$$b_{1(2-3)} = \frac{S\bar{x}_1(x_2 - \bar{x}_2)}{S(x_2 - \bar{x}_2)^2} = \frac{Sx_1x_2 - b_{23\cdot4}\ Sx_1x_3 - b_{24\cdot3}\ Sx_1x_4}{S(x_2 - \bar{x}_2)^2}$$

$$S(x_2 - \bar{x}_2)^2 = Sz^2 = Sx_2^2\ (1 - R^2_{2\cdot34}) = Sx_2^2 - b_{23\cdot4}\ Sx_2x_3 - b_{24\cdot3}\ Sx_2x_4$$

$$\therefore\ S\bar{x}_1^2 = \frac{(Sx_1x_2 - b_{23\cdot4}\ Sx_1x_3 - b_{24\cdot3}\ Sx_1x_4)^2}{Sx_2^2 - b_{23\cdot4}\ Sx_2x_3 - b_{24\cdot3}\ Sx_2x_4}$$

Let i_{12-34} represent the correlation between x_1 (the dependent variable) and $x_2 - \bar{x}_2$.

$$\text{Then } i_{12-34} = \frac{\sigma\bar{x}_1^2}{\sigma x_1^2} = \frac{S\bar{x}_1^2}{Sx_1^2} = \frac{(Sx_1x_2 - b_{23\cdot4}\ Sx_1x_3 - b_{24\cdot3}\ Sx_1x_4)^2}{Sx_1^2(Sx_2^2 - b_{23\cdot4}\ Sx_2x_3 - b_{24\cdot3}\ Sx_2x_4)}$$

Sincerely yours,

T. C. MCCORMICK

COMMITTEE ON GOVERNMENT STATISTICS AND INFORMATION SERVICES

The major staff activities of the Committee on Government Statistics were terminated on June 30, when preliminary reports in the several fields under review were brought to completion. The Committee's skeletonized staff will be engaged for the summer in the preparation of the final report of the Committee and of its affiliate, the Advisory Committee to the Secretary of Labor. This report, to be issued in January, will set forth a summary treatment of the oppor-

tunities for improvement, better planning, and coördination of the statistical services of the Federal Government. Attention necessarily must be given to the statistical activities and needs of the various agencies of the emergency program, as well as to the permanent services. The report will supplement the direct efforts to effect immediate changes and improvements during the year, and will summarize and unify the results of the Committee's active work. In addition to the general report the Committee's memoranda and working materials, insofar as practicable, will be made available to students in the special fields involved.

Since the last note in this JOURNAL (see issue for June, 1934, p. 194) the Committee's activities have been extended into several new fields. Conferences have been held on the statistics of education and of crime and law violation. Mr. Morris Ploscowe of the Harvard Law School was engaged to survey the latter field. Active work in the field of relief and welfare statistics has been carried on during the summer under the direction of Ralph G. Hurlin of the Russell Sage Foundation. Some attention has been given to vital statistics and to certain phases of population statistics. Two interdepartmental conferences were held on industrial classification, and active interest on the part of the agencies concerned is being shown in a proposed general-purpose classification for manufacturing industries. Through the Advisory Committee to the Secretary of Labor attention has been directed to the statistics of industrial accidents and health, immigration, labor turnover, and productivity, in addition to the subjects previously under review. Under the sponsorship of the Bureau of Labor Statistics, the Advisory Committee arranged a meeting of trade-union statisticians on May 18 to ascertain the needs of research workers in this field and to discuss the broad outlines of the Committee's program for labor statistics as thus far developed.

The Committee's representation in the Central Statistical Board is continued under authority of a revised Executive Order (No. 6700, May 4, 1934) which enlarges and more specifically defines the powers of the Board.

MEREDITH B. GIVENS, *Executive Secretary*

July 16, 1934

OUGHT THE UNITED STATES TO RESUME LENDING ABROAD?

A luncheon meeting of the American Statistical Association was held on Wednesday, April 18, 1934, at Busto's Restaurant, 11 Stone Street, New York City. Forty-one persons were present. Mr. Stanley B. Hunt, Economist for the Tubize Chatillon Corporation presided. The following topic was debated: "Ought the United States to Resume Lending Abroad?"

Dr. Marcus Nadler, Director of Research of the International Institute of Finance, opened the case for the affirmative. He contended that exports are essential for the maintenance of prosperity and that the best way to encourage exports is to lend abroad. In his opinion, the choice lies between lending by private individuals, lending by the Government, and direct or indirect subsidy of exports. He considered the proposal to accept silver at a premium of 25 per

cent in payment of exports from the United States, to be primarily an export subsidy. He felt that the subsidy would be particularly bad because the coinage of silver would inflate prices in the United States and bring on all the evils accompanying inflation. Dr. Nadler felt more to be at stake than merely the profit on the goods exported, for he held that stimulation of international trade would tend to pull the world out of depression and that world recovery would bring recovery in the United States.

It was the speaker's contention that the United States Government, by building high tariff walls around the country, has been primarily responsible for the fact that a number of our foreign debtors have not paid their obligations promptly and in full. He considered it remarkable that, under the circumstances, interest on about 60 per cent of private foreign debts to the United States is still being paid as it falls due. He felt that the war debts were political in nature and that their payment in full under present conditions was impossible.

Mr. Lawrence Dennis cross-examined the first speaker. He raised the following questions:

1. When under present circumstances, investments are made in foreign countries, is it expected that the investing nation will ever receive payment or, even the interest? Or is it not necessary for the prosperity of the investing country that foreign loans really constitute a gift in disguise?
2. Can foreigners pay the money borrowed unless the lending country accept an excess of imports over exports equal to the amount to be repaid? And is it possible for an industrialized country like the United States, England, or Germany to be prosperous while receiving a large imports surplus? Therefore can an investing nation afford to collect its foreign loans or the interest thereon?
3. Is it true that, as the first speaker asserted, the Germans have been sincerely anxious to pay their debts? As a matter of fact, was not the desire to escape paying debts one of the chief reasons for the rise of Hitler? Query, therefore, can a debtor nation on international balance afford to pay its debts unless it finds markets for its exports on terms of international exchange which permit the debtor nation to export and maintain a decent standard of living? Can a debtor nation find such markets if all other nations are relying in different degrees on an exports surplus in order to enjoy prosperity?

Mr. John T. Flynn was leader for the negative. He stated that it was not a question of whether we should resume lending and then open our trade walls to foreign goods. We have to recognize that the great majority of Republicans and Democrats are opposed to tariff concessions. At present, both of the leading parties favor a policy of isolation. As long as high tariff barriers are retained, there is little chance that loans made abroad will actually be repaid. When due they will merely be refunded. Under the circumstances, foreign bonds can be sold in the United States only by fooling the bond buyers. These bond buyers would not be investing their money but merely giving it away. He closed by saying:

Foreign lending is a scheme which you cannot fit into your presently popular plan. Of course I know this is not sufficient to keep it out of the plan. Our planning has been truly a wondrous thing to behold.

It can comprehend no end of contradictory and self-canceling schemes.

Our Department of Agriculture says we have 285,000,000 acres planted to food; that to give three square meals a day we ought to have 300,000,000 acres, and that therefore we should eliminate at least a hundred million of our poorest acres from cultivation.

We will pay you, if you are a farmer, not to raise crops. And if you are not a farmer, we will lend you the money to start a farm.

We plan to get about three million farmers off of the farms and about two million city workers out of the towns onto the farms.

We pay farmers millions of dollars to plow under their corn and to raise fewer pigs, then we buy their pigs and feed them into fertilizer machines to sell to farmers to enable them to raise more corn to raise more and better pigs.

We call on the world to disarm and then plan to spend $380,000,000 on battleships. We appropriate money to stimulate the capital goods industry and then bind producers to codes under which they are prevented from increasing their capital investment. We call on business to raise wages and then help business to raise prices to cancel out the wage increase.

Since, therefore, we have plans without blueprints, systems without method, order without arrangement, battles without strategy—in short, planned economy without plans—I do not see why we will not presently arrive at the point when we will call on American investors to resume lending without any provision for payment.

Mr. Maxwell S. Stewart, Economist of the Foreign Policy Association, cross-examined Mr. Flynn. He raised the following questions:

1. Is not the United States at present moving away from the isolationist policies recently pursued?
2. Is not foreign lending so important that it ought to be sponsored by the Federal Government?
3. Has not the difficulty with foreign loans in the past been that they were not handled honestly and intelligently?

To the last inquiry Mr. Flynn countered with the question as to why we would expect future loans to be managed more honestly and intelligently than past loans.

A general discussion from the floor followed. One speaker contended that lending for the purpose of encouraging foreign trade is highly important because every dollar of foreign trade tends to generate many more dollars of domestic trade.

Another speaker raised the question as to whether, if it was legitimate for foreign nations to repudiate their debts on the ground that these debts were political obligations, it was not equally legitimate for the United States Government to repudiate its debts to purchasers of liberty bonds—using the same argument.

WILLFORD I. KING, *Secretary*

TECHNICAL METHODS OF FORECASTING STOCK PRICES

A dinner meeting of the American Statistical Association was held on Tuesday evening, April 24, 1934, at the Roger Smith Restaurant, 41st Street and Madison Avenue, New York City. One hundred and sixty-one persons were in attendance. The general topic for discussion was "Technical Methods of Forecasting Stock Prices."

Donald W. Ellsworth, Editor of the *Annalist*, presided. He initiated the discussion by calling attention to the fact that stock prices are influenced not only by economic forces but also by political forces. The forecaster must take the latter as well as the former into account if he is to be successful. The article by Holbrook Working in this JOURNAL, March, 1933, was referred to as raising the question as to whether or not stock price fluctuations are due merely to chance. Mr. Ellsworth asserted that, in two respects at least, the movements of stock prices are different from cumulative chance phenomenon:

1. Unlike chance data, stock prices cannot go below zero.
2. The fact that stock prices are tied to earnings definitely limits the altitude to which they can climb.

In the discussion which followed the reading of the papers, Mr. Ellsworth asked Frederick R. Macaulay to comment on the relationship between chance phenomena and stock market behavior. In reply, Dr. Macaulay called attention to the fact that the normal curve is only one of the many distributions which may result from chance. If, for example, one threw dice, some of which were loaded but slightly, one would obtain as a result, not a normal but a skew distribution. A chance arrangement of this type might behave in a manner closely resembling the way in which the stock market acts. This would be especially true if, from time to time, the loading was shifted from some of the dice to others. The shorter movements of the stock market behave much like the normal chance distribution. This can scarcely be said, however, of the longer swings.

The first regular speaker of the evening was Mr. H. M. Gartley who presented an outline of methods of technical approach. Mr. Gartley expressed the view that, while forecasters of stock prices should give careful attention to economic and political forces influencing the market, they cannot safely afford to ignore the information contained in the movements of the stock market itself. He stated that the criticism is often made that market changes can do nothing more than record effects, for they come too late to show causes. In his opinion, however, this assumption is incorrect, for, to a great extent, the changes occurring on one day indicate the probable changes which will occur on the next day. The method of utilizing the actions of the market to predict its own future course is not new, having been followed at least since 1885. The Dow Theory is perhaps the best known of the earlier attempts along this line.

The forecaster has two problems. His first is to tell when to buy stocks. Economic data which would be helpful for this purpose often appear too late to be useful. Market data, on the other hand, are available each day that the mar-

ket is open. The second problem of the forecaster is to tell which stocks to buy. The relative behavior of the various stocks in the market offers material assistance in this connection. For example, it is inadvisable to buy a stock the price of which is falling behind the prices of other stocks in the same group.

In closing, Mr. Gartley stated that pending legislation to regulate the stock market, will, if enacted, curtail speculation by an unknown fraction. The chances are that both short selling and speculative manipulation will be lessened. If so, the minor trend movements of the market will probably grow more erratic than in the past and the larger swings of the market will probably be of smaller amplitude than formerly.

Mr. Gartley expressed the opinion that, if a forecasting system became very widely accepted so that it was followed by the majority of speculators, it would probably become useless. A critic from the floor took exception to this point of view and contended that acceptance of a forecast would probably accentuate rather than lessen the predicted movement.

The second speaker of the evening was Mr. R. W. Schabacker, Financial Editor of *Forbes Magazine* who spoke on The Logic of the Technical Approach. Mr. Schabacker pointed out that one of the strong features of the methods used by those who depend primarily upon market movements is that these methods, although frequently giving rise to small losses, rarely result in large losses. The stock market moves in trends. As long as people wish to buy rather than sell, the trend will be upward. When more sellers than buyers appear, the trend will turn downward. It is not important to know why either buyers or sellers are in the majority—all that is necessary is to know that such is the case and align oneself accordingly. As soon as the trend changes, one must also change his position in the market, and the rules of technical analysis make such a change practically automatic.

Mr. Schabacker illustrated his thesis by a few practical examples of chart theory, including the following:

1. When the trend of the market gradually flattens out, a change in direction is likely to occur.
2. A head and shoulders bottom usually indicates that a rising movement is ahead.
3. When a reaction cancels more than two-thirds of the previous advance, the chances are that the decline will go much further.

The third speaker on the program was James F. Hughes, Analyst for Charles D. Barney and Company. His topic was the Intermediate Speculative Approach. Mr. Hughes began by stating that the reason he was more interested in intermediate movements than in long-time swings of the stock market is that the intermediate movements, running from one to six months in length, give to the speculator a sporting chance of obtaining profit from three to four times as large as he could get by following the major movements only. Furthermore, it is easier to forecast intermediate than major movements. This is especially true at present because of the fact that economic forces are materially modified by the New Deal. On the other hand, the market continues to perform as usual,

and the best method of forecasting movements is to use indications from the market itself.

Mr. Hughes emphasized the fact that this type of forecasting is based wholly upon probabilities. It makes no pretense of stating certainties. In some cases, the regularity of performance is great enough to make the probability of successful forecasting high. For example, between 1929 and 1932, the rule held that the summer and winter seasons were marked by advances while the spring and autumn seasons rather regularly showed declines. In nearly every case the technical rally from a temporary bottom amounted to between 45 and 55 per cent of the decline from the last preceding high. Similarly, during the present bull market, each setback has amounted to approximately 50 per cent of the rise just preceding.

In response to a request from the floor, each of the three speakers of the evening registered his forecast of the outlook of the market for the near future. The forecasts are as follows: H. M. Gartley—The market is now at 105 on the Dow-Jones average. The probabilities are that, between now and October 10, 1934, the July, 1933, high will be exceeded materially—prices rising at least to 128. If the October, 1933, low is penetrated before the July, 1933, high has been overtopped, we will be in a bear market. R. W. Schabacker—The market will probably continue for the next six months in a trading range without any sharp advances or declines. The next two months, however, will probably see minor advances. The chances are in favor of a spring rise and a summer decline. Recently, however, the action of the market has been too slight to make it give any very definite forecast. James F. Hughes—The present outlook is very bullish. A peak may be reached around Decoration Day. If, however, the market then pushes ahead, the chances are that the crest will be reached around Labor Day. If the expected spring advance does not materialize, prices will probably run up materially during the summer and reach a summer high about Labor Day.

WILLFORD I. KING, *Secretary*

PROGRESS OF WORK IN THE CENSUS BUREAU

CENSUS BILL LOST IN THE SENATE

The bill to provide for a "census of unemployment, employment, and occupations" to be taken as of November 12, 1934, passed the House on June 7, but was lost in the Senate where it was never brought up for consideration, being left on the calendar when Congress adjourned on June 18. While population was not explicitly included in the scope of the census as defined in the title of the bill, Section 1 provided for the inclusion of "such enumeration of population as the Director of the Census might deem appropriate and necessary for the purposes of this section." That would have given the Director all the authority needed for taking a complete population census. The bill provided that the date of the quinquennial census of agriculture which under the existing law is January 1, 1935, be changed to November 12, 1934, so that it could be taken in conjunction with the census provided for in the bill.

The bill authorized the allocation of $7,540,000 from F.E.R.A. funds to provide for the expenses of this census; also the allocation of $368,465 from the same source to provide for the printing of reports collected and now being compiled on the census of American business, census record preservation, urban tax delinquency, and real property inventory.

The failure to pass this bill therefore has some rather far-reaching consequences. It disposes for the time being of the plan for a population census midway between the decennial censuses. The census of agriculture, which will now be taken in January, 1935, as provided in the existing law, is left with a quite inadequate appropriation of $2,270,000 as provided in the budget of 1934–35. The quinquennial census of agriculture in 1925 cost about $4,400,000. Had the bill passed, a part of the cost of taking the census of agriculture would have been met from funds appropriated for the other census, because in the agricultural districts the canvass for the two censuses would have been conducted by the same enumerators.

Another consequence of the failure of this bill is that there are no funds available for printing the reports of the census of American business, data for which are now being tabulated. It is to be expected that funds for this purpose will ultimately be provided. In the meantime the main results of this census will be published by states, cities, and counties in the form of mimeographed tables as rapidly as the tabulations now in progress are completed. The situation is the same as regards the study of urban tax delinquency and the real property inventory. No funds are available for printing and resort must be had to mimeographed statements.

The "census record preservation" refers to a project for making and binding photostated copies of the population schedules for the censuses of 1800, 1810, and 1820 in order that these copies may be used for reference and the original schedules, which are now becoming worn and dilapidated through constant use, may be preserved. The project also provides for reconditioning and covering with crepeline about 18 volumes of the original schedules. The work is more than three-fourths completed, but cannot now be continued unless the necessary funds are secured from some source.

Since there is to be no census of population, information regarding the farmer's family or the farm population can be obtained only in so far as it may be provided for on the agriculture schedule; and there is under consideration a plan for including on that schedule certain data regarding the members of the farmer's family covering sex, color or race, age, marital condition, whether living on a farm in 1930, occupation followed off the farm if any.

PRINTING DELAYED REPORTS

Through the efforts of the Secretary of Commerce, the Bureau has now been provided with funds for printing the 1931 and 1932 regular annual reports on mortality statistics, birth statistics, and financial statistics of cities. The data for these reports were compiled in regular course as provided by law, but the tabulations though long since completed, have been held in manuscript awaiting time when they could be printed. The funds now secured provide also for

printing the complete report of the 1931 census of manufactures, which will be a bound volume of about 1,300 pages similar in scope and detail to the printed reports of previous censuses of manufactures.

While this is very gratifying and encouraging, the unpleasant fact remains that the printing fund allotted to the Bureau for the current fiscal year is quite inadequate, being only about one-third of what is needed to print the regular reports of the Bureau in their usual size and format.

VITAL STATISTICS

Now that the registration area both for births and for deaths has been completed to cover the entire United States, the Bureau of the Census plans to devote more effort and attention to building up and perfecting the registration. Certain steps have already been taken in that direction:

(1) Conditions in the state of Utah had come to be so bad that it looked as if the state would have to be taken out of the registration area, to which it had been admitted for deaths in 1910 and for births in 1917. The state office was 22 months behind in supplying transcripts of certificates to the Bureau. In the hope of saving the situation the Bureau sent two of its best men to Utah, one being the Assistant Chief Statistician for Vital Statistics, who spent nearly seven months in that state, and with the coöperation of the State Board of Health and the support of the Governor succeeded in effecting a complete reorganization of the state registration work, placing it upon a satisfactory basis.

(2) Under the auspices and with the financial support of the Bureau of the Census a conference of the state registration officials of the southern states, presided over by the Chief Statistician for Vital Statistics, was held in Montgomery, Alabama, in June. This was the first meeting of the kind ever held anywhere and it appears to have been highly successful in awaking interest in registration work and promoting discussion of the best ways of dealing with the problems and difficulties confronting the state registrars. It is planned to hold similar conferences in other sections of the United States as opportunity permits.

(3) An extensive campaign to promote the registration of births is being conducted in the state of Georgia. Through the agency of the Post Office some 500,000 postal cards for reporting births that had occurred within the previous twelve months were distributed to families throughout the state. The returns are being checked to the registered births. While this incidentally furnishes a test of the completeness of the registration, the primary purpose was to educate the public regarding the importance of having births registered, thus protecting the civil and property rights of their children as well as furnishing the basis for satisfactory vital statistics. To give publicity to this campaign use was made of the press, the schools, the radio, and other agencies.

J. A. H.

MISCELLANEOUS NOTES

Changes of Editorship.—Beginning with the December issue of this JOURNAL the present Editor, Frank A. Ross, will be succeeded by Professor Samuel A. Stouffer, of the University of Wisconsin, as Acting Editor. Until February his address is 7028 Commerce Building, Washington, D. C. With the same issue Dr. Ralph J. Watkins, Room 100, Bureau of Business Research, University of Pittsburgh, Pittsburgh, Pennsylvania, will become Acting Review Editor succeeding Dr. Leo Wolman.

Communications regarding the JOURNAL should be sent to these new officers. The business affairs will remain in charge of Dr. Willford I. King, New York University, Secretary-Treasurer of the Association.

The retiring Editors wish to thank the members for their loyal support.

To All Members.—Second-class matter is not forwarded by the Post Office; hence any member changing his address should notify the Secretary-Treasurer promptly in order to assure himself that the JOURNAL will continue to reach him.

The Albany Chapter.—A meeting of the Albany Chapter was held on May 24, 1934, at which thirty members and guests were present. The annual election of officers took place. Dr. R. L. Gillett, of the New York State Department of Agriculture and Markets, was elected President; Dr. W. W. Cox, of the New York State Education Department was elected Vice-President; Dr. David M. Schneider of the New York State Department of Social Welfare was reëlected as Secretary-Treasurer; and Messrs. J. Gibson, of the International Business Machine Corporation and R. F. Bucknam of the New York State Public Service Commission were elected as members of the Executive Committee.

Mr. Robert Whitten of the New York State Planning Board discussed "The Availability and Use of Statistics for State Planning" and Dr. W. W. Cox of the State Education Department delivered a talk entitled "The Job of the Statistician." A lively discussion followed. As a result of this discussion, it was decided that the President of the Chapter appoint a special committee which will go over the Civil Service classifications for statistical positions and which will, with the approval of the Chapter, submit its recommendation to the State Civil Service Department.

Seventh Annual Meeting of the Boston Chapter.—The Seventh Annual Meeting of the Boston Chapter was held at the Boston City Club on Monday evening, May 21, 1934. There was 31 members and guests present. Dinner was served at 6.30. The business session was called to order at 7.30 by the President, Dr. J. Franklin Ebersole. The annual reports of the Secretary-Treasurer were read and approved.

The following officers were elected:

President—Dr. J. Franklin Ebersole, Harvard Graduate School of Business Administration

Vice-President—T. Frederick Brunton, Assistant Actuary, John Hancock Mutual Life Insurance Company

Secretary-Treasurer—Roswell F. Phelps, Director of Statistics, Massachusetts Department of Labor and Industries

Counsellors—J. B. Sedgewick, Security Analyst, Boston; Professor D. S. Tucker, Massachusetts Institute of Technology

The President appointed Messrs. A. Ross Eckler and Roswell F. Phelps to serve as a State Committee to represent the Association in all matters relating to the compilation and use of statistical data in Massachusetts, with authority to add one or more members to the Committee if it should be deemed advisable. The functions of this Committee are: "to initiate measures designed to improve the quality of statistical work conducted by municipal, state and federal agencies, or by private organizations, within the boundaries of the state, to coördinate the activities of the local Chapters and of the National Association, and where desirable, to stimulate the creation of new Chapters."

The following report of the addresses and discussion was prepared by Mr. Joseph B. Hubbard of the Harvard Economic Society.

Mr. J. M. Cassels, of Harvard University, presented a paper on "Certain Aspects of the New England Milk Situation." He emphasized the necessity of following four general principles in making such a statistical investigation: (1) theoretical analysis should be used as a guide to empirical study; (2) an empirical study of the actual institutional arrangements in the industry should be made; (3) data should be carefully appraised when collected; and (4) statistical analysis should be carried out according to the most appropriate methods.

The most important question from the theoretical point of view in the study of the demand for milk in the Boston area was its elasticity. Demand has several sources, since fluid milk may be used directly by the consumer or consumed in such forms as butter, cheese, and ice cream. For the latter products, demand was much more elastic than for fluid milk. The so-called surplus milk of the Boston area is not consumed in fluid form, and the arrangement for collective bargaining shows an institutional adjustment set up in an attempt to solve the problem created by this difference.

For fluid milk, elasticity of demand varied as between wagon delivery and store sales, the latter proving to be the more elastic. Data for the Boston market existed for a sufficiently long period (14 years) to compute normal seasonal change in volume of sales and then study effect of price changes, which had not been frequent in the interval, and could be eliminated in determining the seasonal.

Mr. Einar Jensen, of the Agricultural Adjustment Administration in Boston, in discussing Mr. Cassel's paper, called attention to the fact that the price fixed through the A. A. A. was a producers' price. Class I milk was sold as fluid milk in Boston, and Class II, as surplus milk; they were the same qualities of milk, but sold at different prices according to their use, Class II milk at the lower price. The producer received a composite price made up in the proportion of the two uses to which the entire supply was put.

Mr. Jensen agreed that the demand scale was the basic problem in milk distribution; but questioned whether an instantaneous demand curve answered the question fully. He mentioned the change in consumption habits in San Francisco and Oakland, where an unsatisfactory milk situation and pushing of the consumption of fruit juices by fruit producers had obviously changed demand curves.

Mr. Herman F. Arendtz, in speaking on "Silver," stated that he desired to sweep off some of the cobwebs created by prejudice by advancing certain propositions which he regarded as unassailable. The first of these was that silver must be treated as a monetary metal—only one-quarter of total production was used in the arts. Silver standards go back to early times, while the single gold standard is novel. Silver, unlike paper, has been sound money. Though silver depreciated more than gold since 1900, it did not show the same violent fluctuations. The large credit

structure based on gold, of course, was partly responsible for its wide movements; the point the speaker wished to emphasize was that silver did not fluctuate too much in value.

The speaker emphasized the fact that silver monetization would not call forth floods of silver onto the market, as frequently stated. Summarizing possible supplies from production, from the Orient, and from market stocks and other sources, he estimated that not over 700 million ounces could be drawn out. At the mint price of $1.29 per ounce, the total would amount to only $900,000,000.

As to bimetallism, he pointed out that Gresham's law would result in each country drawing the weaker metal. He regarded this monetary demand as an important stabilizer, and pointed out the stabilizing effect of the Latin Union. He did not believe bimetallism feasible without international agreement.

"Why bother about silver?" was the usual comment on such ideas. In his view, the reason lay in the difficulties encountered in recent years with bank deposit currency. He did not favor a politician managed currency any more than a banker managed currency. He desired less credit and more hard money. His proposal was that the Treasury buy the silver at the market price, not at a fixed ratio, and provide the seller with a silver warehouse receipt which would be redeemed at the price of silver when presented; the proposal contemplated free payment of gold in circulation as well as silver.

In discussing the paper, Professor Ebersole raised the question as to whether our present measures regarding silver might not lead to our sending out gold and so be a force for world recovery or, on the other hand, whether they would merely have the result of deflating the Orient by taking silver from it. Mr. Arendtz saw the possibility that measures relative to silver might break the gold jam, and regarded silver as alone suitable for such a monetary system as he proposed, as against such commodities as manganese, rubber, steel, etc.

The Cleveland Chapter.—On April 2, the Business Statistics Section met to make its semi-annual forecast of the Annalist index of business activity. The final meeting of the section until September was held on June 11. Mrs. Frida F. Selbert of the National Machine Tool Builders' Association discussed the construction and significance of the index of orders of the machine tool industry.

The annual dinner meeting of the Cleveland Chapter was held on Monday evening, May 21, at the Chamber of Commerce Club. Dr. C. E. Gehlke, retiring President, presided. Three papers were presented. D. A. Hill of The Ohio Public Service Corporation was the first speaker. His paper was entitled, "Notes on Some Current Statistical Work," in which he discussed some novel means of recording and tabulating consumption data by consumers. Mr. M. J. Feldstein of the School of Applied Science, Western Reserve University, delivered a paper "On the Measurement of Group Status" in which he discussed the application of statistics in social work. Mr. S. A. Swensrud, Assistant to the President of The Standard Oil Company of Ohio, spoke on "Observations on the Use and Limitations of Statistics in the Operation of Codes."

The following officers were elected for the coming year: Dr. G. E. Harmon of Western Reserve University, President; Mr. L. M. Whittington of The Van Sweringen Company, Vice-President; Mr. E. A. Stephen of The Ohio Bell Telephone Company, Secretary.

The Connecticut Chapter.—In February the Connecticut Chapter held a dinner meeting at which Professor Mark May presented a discussion of "Dangers in

Partial Correlation." At the annual business meeting preceding the dinner Professor R. B. Westerfield of Yale University was elected President for 1934; Professor Burton Camp of Wesleyan was elected Vice-President; and Mr. Paul Dorweiler of the Aetna Life Insurance Company, Hartford, and Mr. Louis Nichols of the General Electric Supply Corporation, Bridgeport, were continued on the Executive Committee.

On June 4 a dinner meeting in New Haven heard a paper on "Size and Earning Power of Corporations" presented by Professor W. L. Crum of Harvard University.

The Pittsburgh Chapter.—Mr. E. C. Molina of the Bell Telephone Laboratories, Inc., New York City, was the guest of the Pittsburgh Chapter at a special luncheon meeting, held April 10 at the Chamber of Commerce. His topic was "Some Comments on Sampling."

At a luncheon meeting held at the Harvard-Yale-Princeton Club, April 26, the subject "The Real Property Inventory" (a CWA project) was discussed by Mr. Theodore A. Veenstra, Bureau of Business Research, University of Pittsburgh.

"The Real Estate Tax in Pittsburgh" was discussed by Dr. J. P. Watson, Assistant Director, Bureau of Business Research, University of Pittsburgh, at the luncheon meeting held at the Harvard-Yale-Princeton Club, May 24.

Mr. Robert A. Burrows, Statistician, A. E. Masten and Company, spoke on "Corporate Financial Analysis" at the luncheon meeting, held June 28 at the Harvard-Yale-Princeton Club. Mr. Burrows is a graduate of the Carnegie Institute of Technology and has had experience as an engineer, teacher, statistician and financial writer.

San Francisco Chapter.—Three-day sessions of the San Francisco Chapter were held on June 20, 21, and 22 in connection with the national meeting of the American Association for the Advancement of Science in Berkeley, this summer.

The first session on June 20th was with the American Mathematical Society, The Econometric Society, and Sections A and K of the AAAS. The general problem was probability and sampling, and was ably covered by four papers. The first by Dr. E. B. Wilson was "Boole's Challenge Problem."[1] The second by Harold Hotelling of Columbia University discussed "Individual Demand Functions with a Limited Budget." John M. Thompson of Rice Institute covered "A Mathematical Theory of Production Stages in Economics." The concluding discussion was by Dunham Jackson of the University of Minnesota. His subject was "Mathematical Principles in the Theory of Small Samples."

The Thursday session was in coöperation with the Econometric Society and Section K, and was devoted to papers on price analysis. Dr. H. E. Erdman of the University of California discussed "A Statistical Analysis of California Walnut Prices," followed by H. J. Stover, also of the Giannini Foundation, on "Some Problems in the Determination of Supply-Price Relationships." The concluding paper was by Holbrook Working of the Food Research Institute of Stanford University. Dr. Working's subject was "Price Differentials and Price Ratios as Subjects of Commodity Price Analysis." There was a prepared discussion on each of these papers, as well as questions from the floor.

The afternoon session on Friday was in coöperation with the Econometric Society and Section K. Norman J. Silberling of San Francisco read a paper on "Commodity Surpluses as Factors in General Price Movements and in Business Fluctuations." This was discussed at length by Joseph S. Davis. The second paper was by John B.

[1] See page 301.

Canning of Stanford University on "A Theorem on a Family of Averages of a Certain Type of Frequency Distribution," followed by Gerhard Tintner, Konjunkturforschungs Institut, Vienna, concerning "Some Notes on the Analysis of Time Series."

A dinner meeting of the same three organizations was held at the Bellevue Hotel in San Francisco on Friday evening. Sixty-five members and guests were present. The first paper was by W. F. C. Nelson of the Cowles Commission for Economic Research. Nelson spoke on "A New Index Number of Investment Experience in American Industrial Common Stocks, 1872–1915." The discussion was a description of the work of the Cowles Commission in constructing an index number of security prices which would adequately reflect national investment experience. The existing index numbers were compared and their defects noted. Most of these were special purpose numbers and not designed to portray investment experience. This paper was followed by Dr. W. L. Crum of Harvard University on "Selected Statistics on the Financial Condition of Corporations." Dr. Crum gave a résumé of his forthcoming paper on this subject. The meeting was declared adjourned at ten o'clock, notwithstanding the many questions which continued to be presented.

The sessions were all well attended and the papers presented were of a high quality. Many of them will be presented in full or abridged form in other publications. Space does not permit more than a listing of their titles at this time.

Dr. E. B. Wilson, former President of the American Statistical Association, addressed a general meeting of the AAAS on Thursday evening on the topic "The Periodogram of American Business Activity. This talk was illustrated with lantern slides, and was very well attended.

United States Bureau of Labor Statistics.—On December 7, 1933, the National Recovery Administrator issued an order "requiring certain statistical reports from members of industries subject to Codes of Fair Competition," and authorized the Bureau of Labor Statistics as his agent "to request, receive, and tabulate such reports." Acting under this order the Bureau has prepared new schedules for employee, pay roll and hour information which is more comprehensive than those used in the past. This order also enables the Bureau to call for 100 per cent coverage on all establishments within any particular industry working under a code. The new schedules are now being used in the woolen and worsted, fertilizer, and paint, varnish, and lacquer manufacturing industries. New industries will be added from time to time as proper arrangements are made with respective code authorities.

The Bureau of Labor Statistics is coöperating with the Consumers Advisory Board under the direction of the Russell Sage Foundation in the making of a study on consumer debts. The Bureau obtained information from 125 companies, representing over 100,000 employees in 22 states and the District of Columbia, with regard to wage executions among their employees. Results of this phase of the broader study will appear in the *Monthly Labor Review*.

The special inquiry into wages and other labor conditions in motor truck, motor bus, air and water transportation, has been completed and the results are now in process of being printed. This inquiry was a joint undertaking of the Bureau of Labor Statistics and the Railroad Coordinator.

The general cost of living study, for which a special appropriation was made at the recent session of Congress, will be started in the late summer or early fall. The appropriation granted was only about half the amount requested and consequently the study will have to be considerably less comprehensive than originally planned.

A survey of unemployment benefit plans in the United States and of unemployment insurance systems in all countries, has recently been made by the Bureau, in supplement to its earlier studies in this field. The present survey deals with conditions in 1933 and 1934. The findings are being published in a series of articles in the *Monthly Labor Review.*

The operation of the various state old age pension systems up to the end of 1933 was described in an article in the August issue of the *Monthly Labor Review.* This article was based on information secured from the several state agencies, and shows the extent to which the state laws had actually been put into operation, the number of pensioners, average amounts of pensions, etc.

The 1933 survey of the Consumers Coöperative Movement in the United States is almost completed. The Bureau has been making periodical surveys of this subject for a number of years, and the findings of the present study will be of particular usefulness because of the increased interest in the coöperative movement in the United States. The Bureau also has in press a short bulletin dealing with the organization of consumers' coöperative organizations, and a similar one dealing with the organization and management of coöperative housing associations is in preparation.

The Bureau is participating very actively in the work of the Committee on Safety and Health Standards for the NRA codes. This Committee was appointed by the Secretary of Labor for the purpose of assisting in the drafting of minimum regulations regarding safety and health for application to the various industries which are operating under NRA codes. Certain general standards for manufacturing industries as a whole have already been prepared, and work is now progressing on individual industries.

On July 1 the Bureau launched a study of wages, hours and working conditions in the petroleum industry. This study is being done in coöperation with the Petroleum Administration and the Planning and Coordinating Committee of the petroleum industry. It covers all parts of the industry, including drilling and production, refining, and both wholesale and retail marketing. The information is being gathered by field agents. It is hoped that this part of the work will be completed by October 1 and the tabulation and analysis by the end of the year.

Women's Bureau, United States Department of Labor.—Because of the increasing numbers of women employed in clerical work, the Women's Bureau has recently completed a study covering 43,000 women office workers in 7 cities—New York, Hartford, Philadelphia, Atlanta, Chicago, Des Moines, and St. Louis. The 314 firms employing these women were chiefly in insurance, banking, investment, mail order, publishing, public utilities, and advertising. Data on employment conditions, policies of management, and certain personal information concerning the women were obtained from personnel records, pay-roll records, and interviews with management. In the general interviews an attempt was made to determine the effects of mechanization on women's employment in office work. Although lack of records of displacement prevented definite conclusions, the report states that the introduction of machines though tending to curb the rapid rise in the number of employees with the increased office functions of modern business, had not displaced any great numbers, the usual practice being to absorb the workers in other departments.

A study of the wages and hours of employed women in Arkansas was made in the winter of 1932–33 at the request of the Commissioner of Labor, who was in charge of relief measures and felt the need of further information on employment and unemployment. Data on earnings and scheduled hours of work were secured for 3,413

white and 546 Negro women at work in 144 establishments in 14 cities and towns. The hours or days worked and the earnings of each woman for a representative pay period were copied from the books. One-fourth of the women included in the study were in the garment industry, the presence of sweatshops from neighboring states being a complicating problem. A second section of the study includes data on the employment experience of 288 women workers in the 12 months ended December 1, 1932, and shows the economic condition of their households.

At the request of the New Jersey Department of Labor an investigation was made in 1932–33 of the hours and earnings of women engaged in retail trade in that state. Later the study was extended into several other states, covering in all 6,923 women employed in 46 department stores in 17 cities. An effort was made, by taking payroll data for a representative week and for a corresponding week some time earlier, to ascertain the trend of employment in the industry and to note changes that had taken place in store hours and methods of payment.

Developments in the Pennsylvania Department of Labor and Industry.—Recent developments in the administration of the Pennsylvania Department of Labor and Industry of particular general interest are the following:

Permanent Unemployment: In an effort to determine whether it is possible to gauge the amount of permanent unemployment and to secure a knowledge of the surpluses and shortages in the various occupations, the Bureau of Employment of the Pennsylvania Department of Labor and Industry is attempting an occupational survey in one county of the state, using the 15,000 application cards on file in the public employment office in that county as the basis for its study.

Vocational Guidance: In another county, through the coöperation of the Foremen's Club, the school board, and the public employment service, a job specification study is being made of all occupations in the county. This information, when compiled, will be used first for the direction of the vocational training program in the county, and second for the purpose of determining whether people in overcrowded occupations can, with little retraining, be fitted for occupations requiring a similar experience.

Special Policing in Industry: Governor Gifford Pinchot has appointed a Commission authorized to look carefully into the practice of using company-paid deputy sheriffs to police industrial towns and properties and especially to investigate the riots and deputy sheriff shooting at Ambridge, Pennsylvania, last October. Members of the Commission are: Shippen Lewis, Philadelphia attorney and authority on labor standards, chairman; John J. Kane, member of State Legislature and city councilman of Pittsburgh; Francis Biddle, Philadelphia attorney; Paul C. Lehman, former judge of Lewistown; and J. W. Madden, University of Pittsburgh law school. J. P. Shalloo, of the Wharton School faculty, University of Pennsylvania, and author of a monograph for the American Academy of Political and Social Science, called, "Private Police, with Special Reference to Pennsylvania," is secretary to the Commission. The investigation promises to disclose for the first time facts about the whole industrial police system in Pennsylvania. It is expected that the report will point the way to a cure of a long festering sore in Pennsylvania industrial relations.

Occupational Disease: The Department is conducting a survey on the incidence of occupational diseases. The survey includes—(1) Physical examinations of 2,800 anthracite coal miners in coöperation with the United States Public Health Service to obtain data on the frequency of miners' asthma. (2) Chemical analysis of samples of air obtained in industrial plants for the purpose of ascertaining conditions as

they relate to poisonous dusts, fumes, and gases. (3) Visitation of employees by registered nurses to determine whether absence from work has been due to sickness arising out of their employment.

Control of Quality in British Industry.—The movement toward the application of statistical method in the control of quality in industry, seems to be making satisfactory progress in England. A pamphlet entitled "The Application of Statistical Methods to Industrial Standardisation and Quality Control," sponsored by the British Standards Institution and put into shape by Dr. E. S. Pearson, is now at the press, while the first Supplement of the Industrial and Agricultural Research Section of the Royal Statistical Society was issued in June, with Part II of the Society's *Journal*.

Fellowships and Grants-in-Aid of the Social Science Research Council.—The Social Science Research Council has announced the award of forty-six Grants-in-Aid of research, fourteen new Fellowship appointments, and one Fellowship reappointment for a period of one year. The Grants-in-Aid total $22,175; the Fellowships, slightly over $48,000.

As in previous years, the Grants-in-Aid were awarded to assist mature scholars in the completion of research projects already well underway.

The Fellowships were designed to afford opportunity for research training, preferably interdisciplinary in nature, rather than to assist in the carrying out of research projects as such. As in previous years, they were open to citizens of the United States and of Canada not over thirty-five years of age who possessed the Ph.D. or its equivalent in research training and experience.

The Council also wishes to call attention to two additional series of training fellowships in the social sciences which will be offered for the first time for 1935–36. One of the new series will be known as Pre-Doctoral Fellowships for Graduate Study, and will be open to citizens of the United States or Canada not over twenty-five years of age who will not have been enrolled in any graduate school for more than one semester before July 1, 1935. Their purpose is to aid exceptionally promising students of the social sciences to obtain research training beginning with the first year of graduate study. The other new series will be known as Pre-Doctoral Field Fellowships, and will be open to citizens of the United States or Canada not over twenty-seven years of age who are candidates for the Ph.D. degree, and who will have completed prior to the end of the academic year 1934–35 all courses and examinations for which they are eligible before completion of the thesis. The purpose of these Field Fellowships is to supplement formal graduate study by opportunities for field work which will assure first-hand familiarity with the data of social science in the making. The Post-Doctoral Training Fellowships will again be offered under policies and regulations similar to those previously in effect.

The closing date for the receipt of applications for Grants-in-Aid for the academic year 1935–36 will be January 15, 1935; for Pre-Doctoral Field Fellowships and Post-Doctoral Training Fellowships, December 1, 1934; for Pre-Doctoral Fellowships for Graduate Study, December 15, 1934. In order to facilitate the filing of applications on the proper blanks before the closing dates, it is requested that persons interested communicate with the Secretary for Fellowships and Grants-in-Aid, 230 Park Avenue, New York, New York, as early in the fall of 1934 as possible. The first letter of inquiry should include a brief statement of the candidate's proposed plan of work and of his academic and professional record.

PERSONALS

Dr. Charles Josiah Galpin retired June 30 from his post as head of the Division of Farm Population and Rural Life in the Bureau of Agricultural Economics, Department of Agriculture, after fifteen years with the Department and twenty-five years of leadership in the field of rural social planning. During this period he has attained world-wide reputation as a leader in his field and has been tendered special recognition by two foreign governments. He was educated at Colgate, Harvard, and Wisconsin Universities, receiving the degree Litt.D. from Colgate in 1919.

The Board of Directors of the Metropolitan Life Insurance Company on June 26 appointed Dr. Alfred J. Lotka Assistant Statistician of the Company. Dr. Lotka is an active member of the Association, having contributed frequently to this JOURNAL, and has otherwise been of signal service.

Dr. Claudius T. Murchison was appointed in June as Director of the Bureau of Foreign and Domestic Commerce, Department of Commerce.

Dr. J. Neyman of the Biometric Laboratory, Nencki Institute, Warsaw, has been spending the summer in the Department of Applied Statistics, University College, London, where he has been lecturing on recent advances in the application of probability theory to the solution of problems of statistical estimation and the testing of hypotheses. It is probable that he will continue these lectures in London during the Session 1934–35.

Mr. E. S. Macphail, for over thirty-three years connected with the statistical service of Canada and for the last sixteen years the head of the Demography Branch of the Dominion Bureau of Statistics, was superannuated in July, after he had taken a prominent part in the administration of four decennial censuses of the Dominion and three quinquennial censuses of the Prairie Provinces. Mr. Macphail, who is well known to many members of the American Statistical Association through his official position as head of the census and vital statistics of Canada, has made a distinguished contribution to the great improvement of Canadian statistics in both these fields in recent years. Mr. Macphail is a Fellow of the American Public Health Association and a Member of the Council of its Vital Statistics Section. He was the prime mover in the organization of the Vital Statistics Section of the Canadian Public Health Association and is a member of the Section Council. He has contributed valuable articles to the Journals of both Associations.

F. L. Carmichael, professor of statistics and director of business studies at the University of Denver, has received an appointment as statistical consultant with the Petroleum Administrative Board in Washington. Professor Carmichael assumed his duties in Washington July 23.

Dr. Kenneth W. Rowe, former professor at Ottawa College, Ottawa, Kansas, has been added to the faculty of the University of Denver as assistant professor in the department of economics.

Rolf Nugent, who for several years has been on the staff of the Department of Remedial Loans of the Russell Sage Foundation, has been appointed Associate Director of the Department, and is Acting Director of the Department in the absence of Leon Henderson, who is serving as Director of Research and Planning of the NRA. As part of the program of this Department, Mr. Nugent is now conducting a survey of delinquent consumer debt in the United States for the Committee on Delinquent Consumer Debt of the Consumer's Advisory Board of the NRA.

MEMBERS ADDED SINCE JUNE, 1934

Altschul, Professor Eugen, Guest-professor, University of Minnesota, Minneapolis, Minnesota; Associate of the National Bureau of Economic Research, New York City

Anderson, Hugo A., Real Estate and Insurance, 462 Bergen Street, Brooklyn, New York

Beckman, Dr. Theodore N., Professor of Marketing, Ohio State University, Columbus, Ohio; Consulting Expert, Bureau of Census, Department of Commerce, Washington, D. C.

Blum, Natalie, Research, Federal Coördinator of Transportation, 60 Hudson Street, New York City

Boatwright, Dr. John W., Sales Research Department, Standard Oil Company, 910 South Michigan Avenue, Chicago, Illinois

Bratt, Irving L., Student, New York University, School of Commerce, Washington Square East, New York City

Callender, Everett, Scudder, Stevens and Clark, 1 Wall Street, New York City

Carruthers, Lemuel A., Bureau of the Census, Department of Commerce, Washington, D. C.

Cheney, Roy A., Managing Director, Underwear Institute, 528 Union Trust Building, Washington, D. C.

Crowder, Edward T., Jr., Assistant, Department of Economics, Washington Square College, New York University, Washington Square East, New York City

Dunbar, Agnes, National Board Y.W.C.A., 600 Lexington Avenue, New York City

Ferguson, Leonard W., Psychology Department, Stanford University, California

Fox, Philip G., Associate Professor of Business Administration, University of Wisconsin, Madison, Wisconsin

Galloway, Dr. George B., Research and Planning Division, National Recovery Administration, 3420 Commerce Building, Washington, D. C.

Hamilton, Dr. C. Horace, Teaching and research, State College of Agriculture, University of North Carolina, Raleigh, North Carolina

Holland, Reuben, 19 East 47 Street, New York City

Hopkins, John W., Division of Biology and Agriculture, National Research Laboratories, Ottawa, Canada

Hunter, Kenneth H., Retail Price Project, Bureau of Labor Statistics, U. S. Department of Labor, Washington, D. C.

Jaffe, A. J., 4402 South State Street, Chicago, Illinois

Kawakami, Dr. Riiti, Rockefeller Foundation, 49 West 49 Street, New York City

Kennedy, John L., Student, Stanford University, California

Kerschbaum, Paul R., Division of Statistics and Research, Department of Labor, Chicago, Illinois

Knott, Mrs. Grace W., Division of Research and Planning, National Recovery Administration, Washington, D. C.

Kuhns, William R., Managing Editor, American Bankers Association Journal, 22 East 40 Street, New York City

Lemal, Richard F., New York Telephone Company, 140 West Street, New York City

Lindman, Bertram H., Engineering Economist, Highway Cost Commission, Olympia, Washington

Lipkowitz, Irving, 195–15 Central Avenue, St. Albans, Long Island, New York
Lipkowitz, Samuel, 133 Lott Street, Brooklyn, New York
Loomis, Dr. Alice M., Yale Institute of Human Relations, 333 Cedar Street, New Haven, Connecticut
Mackenzie, K. H., Statistical Department, 903 Federal Reserve Bank of Cleveland, Ohio
MacRoberts, D. T., United Gas System, Houston, Texas
McCaw, William, Howard Whipple Green and Associates, 1900 Euclid Avenue, Cleveland, Ohio
Moss, Maude F., Jewish Board of Guardians, 228 East 19 Street, New York City
Nathan, Rita J., Student, New York University, School of Commerce, Washington Square East, New York City
Newman, Leslie H., Jr., Personnel Department, Brooklyn Edison Gas Company, Inc., 380 Pearl Street, Brooklyn, New York
Raskin, George S., Student, New York University, School of Commerce, Washington Square East, New York City
Rich, Maurice H., Auditor, 15 West 36 Street, New York City
Robison, Jesse, Halle and Stieglitz, Room 1400, 25 Broad Street, New York City
Rogers, Roger R., Prudential Insurance Company, Newark, New Jersey
Salman, Bernard, Neuhut, Mansbach and Plonn, 60 Broad Street, New York City
Sayler, Allen W., Federal Emergency Relief Administration, 1734 New York Avenue, N.W., Washington, D. C.
Schneider, Margaret G., Twentieth Century Fund, 11 West 42 Street, New York City
Schwartz, Benjamin C., 1275 Nelson Avenue, New York City
Shipley, L. P., Investment Management Department, Brown Brothers, Harriman and Company, 59 Wall Street, New York City
Southworth, Constant, Consumers Advisory Board, NRA, Washington, D. C.
Stolz, Regina, Department of Statistics, Russell Sage Foundation, 130 East 22 Street, New York City
Timm, Walter E., E. I. DuPont de Nemours and Company, 350 Fifth Avenue, New York City
Vaughn, S. Marie, Citizens' Committee on Relief and Employment, 2221 Locust Avenue, St. Louis, Missouri
Watkins, Dr. J. H., Instructor, Yale School of Medicine, Department of Public Health, 310 Cedar Street, New Haven, Connecticut
Winterbottom, Edmund, Consolidated Gas Company of New York, 310 East Kingsbridge Road, New York City

REVIEWS

Twenty Years of Federal Reserve Policy, by S. E. Harris. Cambridge: Harvard University Press. 1933. Two Volumes: 431, xxxix; 435–865, xi pp.

Twenty Years of Federal Reserve Policy is a detailed account of the workings of the Federal Reserve System from its inception down to the end of 1932, or a little beyond that date. There have been several schools of discussion and writing about Federal Reserve affairs; the first, political and "inspirational"; the second, descriptive and analytical; and the third, theoretical and general. Mr. Harris' volume belongs in the second group. It is the result of extensive investigation and comparison of available sources of information with elaborate footnotes intended to bear out the findings at which the author arrives. The work is, in many respects, meticulously done; and brings together useful collections of data in a convenient form.

Mr. Harris' title is, however, a misnomer; his book contains relatively little of discussion of Federal Reserve policies although much of Federal Reserve annals and operations. The style of presentation is, perhaps necessarily, dry, with detailed presentation of facts and incidents that have only an unimportant bearing upon the subject under discussion. This naturally prevents the author from seeing the larger aspects of the subject he is discussing, and obviously prohibits him from drawing the lessons from the often heterogeneous masses of data he has gathered, that might otherwise be derived from them. Although he furnishes here and there, a short chapter of "conclusions," his conclusions are usually inconclusive, and their meaning and point frequently hard to infer. This gives the volume an undigested appearance, as if the writer had been overwhelmed by his material and had not been able to form a satisfactory framework within which to arrange it. For these reasons, the work of Mr. Harris, although carefully and conservatively stated, does not make the impression upon the reader's mind that should be expected from it; and it frequently happens that the principal value of one of his chapters is found in something that has incidentally been stated, but apparently has been given but little weight by the author. The volumes are thus, as already said, really annals of the Reserve System rather than a review, much less a critique, of policies.

Mr. Harris has added to his work on the Reserve System itself, appendices in which he gives charts and bibliographies that will be of value. Even these lengthy compilations, however, are in some ways curiously lacking in point, and it is especially to be regretted that the author has included the "addendum" at the end of Volume 2 in which he discusses the banking crisis of 1933. At the opening of this addendum he undertakes to "relate what follows, to the trends of earlier years discussed in this book," but it is difficult to note any such satisfactory relation. The addendum was evidently written with inadequate data at hand, indeed could hardly have been otherwise written at the actual

date of its publication. It consequently reaches, again, only unsatisfactory and indecisive "conclusions," some of which have unavoidably been proven inadequate by subsequent events. It would have been considerably better had the volume been limited to its original scope, as a detailed survey of Reserve operations year by year.

From what has been said it will be evident that, in the opinion of the present reviewer, Mr. Harris' laborious work falls between two stools. It is neither sufficiently "historical" and sufficiently "political" to enable the reader to gather from it the reasons why given Federal Reserve policies were initiated, and then to follow these policies through to their results; nor is it sufficiently "practical" and technical to render it a critique of what has actually been done and its financial effects in practice. At its best it thus furnishes material to which inquirers may be referred, or upon which they may draw, in their study of the subject. In all central banking discussion in the United States during the past thirty years, the trouble has been that economic writers have been too far behind the current of events. They have been unable or unwilling to commit themselves as to the meaning of given tendencies or phenomena, and thus the time has passed when their analyses and their forecasts were likely to be of much service. Even while Mr. Harris' book, for example, was in process of publication, fundamental changes were occurring in our banking system which have shown the necessity of an emphasis entirely different than the one which he has furnished in his volumes—so far as he gives any.

A more effective type of central and Reserve banking discussion is needed in the United States, and until it is furnished the system is likely to go its way as it has up to date, experimenting in all sorts of impossible directions and making the most egregious blunders, which are then solemnly analyzed and documented by economic and financial writers. These remarks in no way reflect upon the character of Mr. Harris' work as such. There is every reason to believe it painstaking and conscientious, probably laying the foundation for a much more mature production by him as its outgrowth. It will be of service along the lines already indicated, but the needs of central banking discussion in the United States prescribe a different sort of analysis, and call for more forward-looking appreciation of problems.

H. Parker Willis

Whither Latin America? An Introduction to Its Economic and Social Problems, by Frank Tannenbaum. New York: Thomas Y. Crowell Company. 1934. xix, 185 pp.

Mr. Tannenbaum has written an effective book. *Whither Latin America* is an economic survey of the countries—island and mainland—to the south of the United States including in its scope such topics as population, industrial possibilities, finance, foreign trade, transportation, education, labor and agriculture. The significance of the book goes beyond the immediate material that is presented, for it demonstrates how loose much of the economic thought on the undeveloped parts of the world has been. In viewing backward countries, it is

almost traditional to assume that progress for them must follow a set pattern and that in the course of time they will experience the same sort of industrialization that has occurred in Great Britain, Germany and the United States. The author questions the validity of any such assumption particularly in its application to Latin America. In an early chapter on industrialism, Mr. Tannenbaum points out that the resources of the continent, especially the reserves of coal, are quite inadequate for the development of any heavy manufacturing and that it will even be necessary to import the equipment for other types of industry.

In a later chapter on transport, the author demolishes the accepted concept that prosperity and progress will follow the construction of railroads. In Latin America, there are no materials for the building of railroads and railway equipment, or fuel to run them. There is little coal and few industrial products to contribute to a profitable traffic and in most areas the population is too scanty. Only in a few sections where agricultural products or minerals in quantity must be moved to the seacoast has the construction of railroads been justifiable economically. Examples are given of existing railroads burdened with the costs of imported supplies and equipment and with operating and maintenance costs so high that goods can be carried in competition more cheaply by llama pack trains. It is probable that such railroads should be abandoned as luxuries too expensive for governments whose budgets are persistently out of balance. In place of any expectation of an inevitable development of transport resulting in progress, it should be appreciated that certain areas of Latin America must be content indefinitely with pack and animal transport on trails and that backwardness and isolation are to be the permanent condition of much of the southern continent. It is quite evident that for any sound development of transport and industry a thorough knowledge of economic geography is essential if the means of communication—rail, highway and water—are to be properly coördinated and if the manufacturing industries are to have any relation to the available resources.

The somewhat disillusioning analysis continues with an examination of foreign trade. It is pointed out that the export trade of Latin America rests upon agricultural products and a few minerals and that for the most of the countries the really important exports are limited to one or two commodities—nitrates and copper for Chili, tin for Bolivia, wheat and meat for the Argentine, coffee for Brazil, coffee and bananas for Central America, sugar for Cuba. It is upon these exports that the present economic advance of the countries has been built. They have paid for the railroads, in so far as the railroads have been paid for, and for such industrialization as has occurred.

For the future, Mr. Tannenbaum paints a most gloomy picture. He sees no chance for any material increase in the export of agricultural commodities with a stationary or declining population in Europe and he is of the opinion that there may even be a decline if nationalism persists and competing sources of supply continue to develop.

Whither Latin America is not a closely documented and factual research project, but rather a survey of problems that should be investigated. In the

absence of facts, the author refuses to state any conclusions explicitly, but he employs the ingenious device of raising questions that demand further study. His questions are so phrased that there is never any doubt regarding the author's opinions or his thesis. They convey a more positive impression than would dogmatic assertions.

It should be stated that with few exceptions Mr. Tannenbaum's questions show no bias though there is frequently room for differences of opinion. For example, it would appear that industrialization and the iron and steel industry are to the author almost synonymous. It is true that Latin America does lack the resources for a primary steel industry, but there are many other forms of manufacturing that could be developed. Indeed, some of the author's questions with regard to foreign trade and agriculture would suggest that industrialization of a kind is inevitable. If there is little opportunity for the developing of new agricultural products and if the markets for the existing products are to be curtailed, the commodity balance in foreign trade may easily become a problem. An examination of the import trade of the Argentine, Brazil or almost any other Latin American country shows a number of items—cotton piece goods, yarn, hosiery, clothing, leather, pottery and other light consumers goods—that might be fabricated at home without any great dependence on iron and steel or power. In this respect, South America is not so much worse off than the countries of Asia or of other unindustrialized parts of the world, for if industrialization is to coincide with supplies of iron ore and coal, it would seem that the world's industrial areas must remain confined pretty closely to the shores of the north Atlantic.

It is doubtful whether the author is justified in his assumption that literacy and widespread education are possible only in industrial nations. Literacy is exceptionally high in Iceland, Norway, Sweden and Denmark—all agricultural or pastoral countries, and education was well advanced in the United States before the industrial era. In Asia, there is the example of Japan which, despite outward appearances, is still essentially an agricultural country.

It would seem also that the author is somewhat too hasty and too pessimistic in his belief that there is little room for diversifying agricultural crops in Latin America or for increasing some of the existing crops. His arguments have a logical sound, particularly in the present period of depression, but they might be applied with equal force to the economic position and foreign trade of almost any country, manufacturing or agricultural. They are valid only if the expansion of world trade is permanently at an end and if all countries are to exist in self-sufficient, nationalistic compartments.

Though *Whither Latin America* is a stimulating survey, it is by no means the final economic word on the southern continent. It is full of ideas and it is stimulating, but it has few facts. It calls for the discarding of traditional thinking and for the examination of Latin America in terms of the conditions found there and not as another Europe or United States. It maps a course for investigation that other students may follow with profit, but it leaves to them the painstaking and prosaic task of grubbing for facts, the searching through official reports for

fugitive data, the visits to factories and plantations, the interminable interviews, the weighing and the analyzing of a mass of material. It is possible that when all this basic work has been done, the answers to Mr. Tannenbaum's questions will not be those implied in their phrasing.

JOHN E. ORCHARD

Columbia University

Ten Thousand Out of Work, by Ewan Clague and Webster Powell. Philadelphia: University of Pennsylvania Press. 1933. 188, xvii pp.

This volume is a study to determine what are the distinctive personal characteristics and socio-economic circumstances of the unemployed. Approximately ten thousand applicants for work relief in Philadelphia in 1931 were interviewed regarding personal characteristics, composition of the family, number of wage-earners per family, occupational distribution, duration of longest job held, length of period of unemployment, and means of "getting along" during unemployment. Fourteen hundred and thirty-nine cases were studied with special reference to family economics. On the basis of the statistical results of these inquiries conclusions are drawn as to the education, wages, and stability of workers and the interrelationship of these factors; the responsibility of the worker and of industry for the predicament of unemployment; and the ways and means of preventing or providing for destitution.

Among the more interesting findings of the study is the relatively high degree of employment stability of the cases. More than one-fourth of the entire working life (years over 18) of these men had been spent on a single job, i.e. with the same employer. There is no evidence in this or in other respects that the group studied represented marginal workers, but on the contrary that they were rather representative of industrial workers in general.

The six appendices contain a careful and rather complete account of the methods employed especially as regards the process of sampling and the reliability of the data. The discussion of the latter is especially thorough. Perhaps too much is made of the high degree of stability in answers shown by the cases inadvertently interviewed in duplicate at varied intervals of time, as their consistency may merely reflect the consistency of error. To the extent that the uncorroborated testimony of persons interviewed is the basis of the information secured, it is of course subject to the usual errors of such data. The study as a whole however, is methodologically very superior to most inquiries of this kind that have appeared, and its conclusions are correspondingly significant.

GEORGE A. LUNDBERG

Columbia University

An Introduction to Statistical Analysis, by C. H. Richardson. New York: Harcourt, Brace and Company. 1934. xi, 285 pp.

The material of Professor Richardson's book is presented under the following eleven chapter headings: Introduction; Tabular and Graphical Representa-

tion: Frequency Distributions; Measures of Central Tendency; Measurement of Dispersion; Skewness: Excess: Moments; Linear Trends; Simple Correlation; Nonlinear Trends: Curve-Fitting; Permutations, Combinations and Probability; The Point Binominal and the Normal Curve; The Theory of Sampling: Measures of Reliability. The appendix contains a selected list of books for supplementary reading and a table of areas and ordinates of the normal curve. An index is given on page 283.

Both author and publisher claim that the student who has but comparatively little training in mathematics can comprehend the text. The writer believes, however, that the student will find considerable difficulty in profiting from Professor Richardson's book without guidance from a competent teacher.

In Chapter 1 the author discusses in detail "Sums and Summations," "Decimal Accuracy," "Significant Figures," "Rounding off Numbers," "Errors in Calculations," "The Propagation of Errors." Technically speaking these phases are details of great importance and well brought out in the discussion, but the student will find this approach confusing and will doubtless welcome the refined points of technique at some more advanced stage of the text or in an appendix.

In Chapter 2, entitled "Tabular and Graphical Representation: Frequency Distributions," a detailed description relative to the preparation of various types of tables should have been included. With the exception of a few simple tables, practically all of the space is devoted to frequency tables. The various forms of graphic presentation are limited and not brought out sufficiently.

A great number of illustrations in the text (the writer counted 25) relate to some form of frequency distribution. No illustration is shown pertaining to semi-logarithmic presentation of time series—the author prefers the words temporal distributions. The important use of the log—log grid is not included in the illustrations nor is a scatter diagram given showing the functional relationship of two phenomena in order to clarify the concept of correlation.

The index does not contain a reference to seasonal variation nor are the methods shown by which time series can be adjusted for seasonal variation. A discussion of the concept of simple index numbers, their preparation and uses in the statistical field, has not been included. This is explained in the preface on the grounds that "No attempt has been made to make this text an exhaustive treatise on statistical analysis. Many topics, such as multiple correlation and frequency curves, have been studiously omitted."

"We have tried to keep in mind that we are writing an introduction that would include the minimum essentials, at the same time hoping this introduction might inspire the reader to continue his study into the more advanced fields."

Although all mathematical processes presented show the training and experience of the skilled mathematician, yet the treatise lacks that desirable unity of presentation which is so important to the beginner and which apparently is so difficult for most authors to accomplish.

The announcement of the publishing house places Professor Richardson's

book as "the best in a decade." This superlative on the part of the publisher is open to criticism. During the last decade numerous American statisticians have made real contributions to the literature of statistical textbooks. Among the outstanding texts of the last decade which deal with statistical methods and which are suited for the beginner are those of Frederick C. Mills, R. E. Chaddock and Frisbee and Riggleman. These books are now internationally known and must be considered as belonging to the best American texts on the subject of statistical analysis.

R. VON HUHN

Health and Environment, by Edgar Sydenstricker. New York: McGraw-Hill Book Company. 1933. 217 pp.

This work is one of the monographs basic to the section on Health in *Recent Social Trends*. The title of the work is most pretentious but Sydenstricker in his preface points out that "the really significant knowledge . . . is not large and can be put in a few pages," provided one sticks to the relevant facts and a reasonable interpretation.

The introduction presents a short but comprehensive survey of two health determining factors: heredity and environment. Speaking briefly of the works of the English biometricians, environment is treated with the recognition that it can only affect individuals who are already biologically diverse. Environment must be considered in its physical, social and biological aspects, each of which has its own peculiar effect on health. Health—or the healthy state—is defined as "freedom from physical and mental impairments (except those that are the natural concomitants of the aging process, the wearing out of the human machine), from disease, and from illness" (p. 18). To measure the state of health for a people we ought to have some kind of an index, but in our present situation such an index could scarcely be more than a "device with which one may divine the obvious."

After a survey of the "Nature and Extent of Ill Health" in which physical impairments, mental impairments, illness and mortality are discussed, a chapter is given to the effect of each of a number of phases of environment on health —geographic, urban and rural, economic status, social, and occupational.

An example of a geographic effect is the prevalence in certain areas of goitre, due to the lack of iodine in the drinking water. The importance of the geographic factor has not been thoroughly studied, and its part as a cause may be misinterpreted. For instance, pellagra which almost invariably appears in the south can easily mislead a student unless joint causes are studied. Our registration has been too small until the last few years to give much analytical help. Conversation with a public health statistician today emphasizes the author's contention that too little study has been made of available data, and among the studies completed too little attention has been paid to joint causes. The differences in the incidence of disease and illness in urban and rural sections has long been recognized. It is interesting that Sydenstricker has had to go back

to 1910, a period before the development of hospitalization, in order to study comparable mortality data in country and city. Our habit of recording deaths at the place of death and births at the place of birth has invalidated the use of statistics for the differential interpretations.

The effect of economic status is that which we would suspect—"the lower the income, the higher the illness rate" (p. 87). Recourse for these results has been taken to a few studies made in industrial centers in the United States in the last several years, and to English data for the effect of social classes on health. Many illustrations are used for showing the effect of social environment on health.

Occupational environment has had its serious noticeable effects on the male population particularly. It must be recognized, however, that occupation cannot be separated from economic condition and the urban situation in which it is usually found.

In a summary chapter the trend of mortality is discussed. It is pointed out that the length of life is still "three score years and ten" with "no evidence that the span of life has increased since ancient times" (p. 161). The expectation of life at different ages, however, has fluctuated. The greater decrease in infant mortality in the last few years has had an important effect on life expectancy at the early ages. A discussion of the trends in various diseases as causes of death is of interest. More interpretation would be helpful at this point. For a number of years the death rate from appendicitis has continued upward. An explanation of this situation does not appear although it would likely be given in terms of joint causes or more probably to attendant secondary or tertiary conditions which have become more dangerous in an aging population.

Sydenstricker has of necessity drawn heavily on studies made under the U. S. Public Health Service and the Milbank Foundation. The major contribution of *Health and Environment* to faithful followers of the work of these two groups is Sydenstricker's assembly and organization of scattered materials.

Relatively few noticeable errors have crept into this book which contains many tables and graphs. On page 62 a reference is made to page 35 instead of page 26. In table 19 on page 78 the death rate per 1,000 of rural males aged 65–74 should be 43.6 instead of 143.6 if the ratio of urban to rural is correct at 1.3.

LESTER S. KELLOGG

University of Buffalo

Regulating an Industry: The Rhenish-Westphalian Coal Syndicate, 1893–1929, by Archibald H. Stockder. New York: Columbia University Press. 1932. 159 pp.

In this period of cartelization under the NRA, it is enlightening to know in some detail what the result of a similar policy has been abroad. The Rhenish-Westphalian coal producers have had many years of experience with collective action, which of course has been unhampered in Germany by anti-trust laws.

Indeed, government compulsion as a buttress of business combination in this industry was employed as early as 1915. Mr. Stockder tells the story of the organization with admirable precision and condensation.

The lesson for us, however, is not entirely clear on the basis of this history of the experience. Before the War the effort to regularize production and maintain prices was continually imperiled by the usual difficulties of voluntary combinations. Agreements either were not fulfilled or placed such unequal burdens on various parts of the industry that they were always being upset and revised. Attempts to control the output of coal continually ran athwart the fluctuations of the demand, which created new and unforeseen problems. Counter-combinations of wholesale and retail dealers and of ultimate consumers were formed to combat the Syndicate's control over prices. Finally, on account of the growth of integrated steel companies providing their own coal, the Syndicate was on the point of dissolution when the War broke out. Governmental regulation during abnormal war conditions is scarcely a parallel to our present situation, and since then Germany has been through a whole series of crises—the abortive attempt at "socialization" in 1919, the inflation, the attempt to pay reparations, and the period of living on loans from abroad, all of which must have given any single institution such as this numerous incalculable biases. With such modifying influences Mr. Stockder hardly reckons.

He concludes, in general, that the Syndicate was finally held together only by governmental pressure, that it had if anything a retarding influence on integration, and that it did not limit the formation of new companies. Control of production was dependent on the market, and production quotas were more an instrument of apportioning orders than of limiting output. But they penalized the more efficient and discouraged technical improvement—except in the case of companies which produced for their own consumption. Seasonal variations in output were greatly reduced, and prior to 1914 cyclical variations were not prominent. On these accounts he believes the regulation of production worked successfully.

Market policy was difficult and often unsatisfactory. Mr. Stockder believes, however, that price policy has been moderate and fair to consumers except in 1900–1901, but this conservatism he attributes largely to the competition furnished by other fuels and sources of energy. Unemployment has been reduced but not eliminated, wages have absorbed a larger share of the product since the War, and social insurance imposed by law has helped to keep labor satisfied—though before 1919 collective bargaining was unknown in the industry.

It will be noted, however, that whatever benefits were achieved by the combination were largely the result of throwing burdens and risks on other parts of the economic complex. Regularization of production was the result of forcing dealers to store stocks, and maintenance of prices largely the result of dumping surpluses abroad. There is no evidence here of brilliant success to be achieved in the stabilization of separate industries, if the success be measured by general welfare.

GEORGE SOULE

Scientific Management: An Aid to Industrial Control, by George Filipetti. Minneapolis: University of Minnesota Press. 1933. 58 pp.

This monograph is a plea for better industrial management on the part of individual concerns and for the interchange of statistical information concerning plant operations. The factual materials on which the discussion rests were obtained by the Employment Stabilization Research Institute of the University of Minnesota, in a survey of manufacturing enterprises operating in the Twin Cities and Duluth.

The manuscript was prepared while the National Industrial Recovery Act was before Congress. The author regarded this as significant, saying that the principles set forth in the study "are essential to the successful application of such legislation . . . this bulletin attempts to point the way to better regulated production and *cost reduction* rather than to restricted production and price maintenance" (pp. 5–6, italics the author's). Whether industry has followed the paths indicated may be left to the reader to decide.

The test of good management is declared to be neither satisfactory sales nor high profits, but whether "an industry can *pay the market prices*" for the available agents of production and then "*manufacture and sell its product at a low cost*" (p. 15, italics the author's). No definition or criterion of "low cost" is developed, but the author finds *prima facie* evidence of inefficient production in the facts that inadequate statistical records are kept by most companies, that few have developed adequate cost accounting systems, production control systems or performance standards, and that out "of a total of 221 companies answering an inquiry as to whether they budgeted and planned production, 162, or 73 per cent, answered negatively" (p. 28).

The monograph is well written, and the analyses are carefully presented. It might have been well to supply a summary table showing the complete situation for all the companies surveyed; the reader gathers that the number of plants answering certain questions ranges from 95 to perhaps 244, but the number of companies covered by the inquiry as a whole seems nowhere to be stated. Possibly it was felt best not to give the figure because many companies answered only a few of the questions asked.

But this is a minor defect. The important, if depressing, contribution made by the bulletin is to show (taking the Minnesota industrial area as a sample) the small proportion of enterprises that habitually plan their own activities in any thoroughgoing fashion (only 59 out of 221 make production estimates for the year ahead) and, by implication, to raise this question: how good are the long-run prospects for successful national industrial planning when relatively few managements as yet keep adequate statistical records or are acquainted with the technique of controlling individual operations even within the plant?

Discussion of this point would carry us well beyond the limits of either this review or the author's study. But for Professor Filipetti to raise the question is surely legitimate, and his monograph provides much food for thought in this connection.

RALPH C. EPSTEIN

University of Buffalo

Nuova Collana di Economisti Stranieri e Italiani, edited by Giorgio Mortara. Vol. VI: *Cicli Economici*. Torino: Unione Tipografico-Editrice Torinese. 1932. 634 pp.

This volume, the sixth according to the formal plan of the important series the inauguration of which was noted by the present reviewer in the April, 1934, issue of the *Journal of Political Economy*—though it is the third volume of that series in order of actual publication—continues to honor the canon, adopted with such admirable results in the two preceding volumes, that the science of economics is not to be regarded as the exclusive property of any single linguistic group, but must continue to grow upon the basis of contributions made by representatives of every one of the great national traditions. In the volume under review, the subject of which is *Economic Cycles*, the English-speaking group is represented by a translation of the introductory chapter of Professor Wesley Mitchell's *Business Cycles: The Problem and its Setting;* the German-speaking group by a translation of Dr. Ernst Wagemann's smaller volume on *Konjunkturlehre;* while the Italian literature is represented not only by Professor Mortara's own interesting introduction to the volume under review, but also by a contribution from Professor Bresciani-Turroni, who presents here a reworking of the able studies previously published by him on the subject to which the volume is dedicated.

It would, of course, have been impossible, in a work of this kind, to omit an example of the work of Wesley Mitchell. If one has any regret with respect to the inclusion of a part of Professor Mitchell's book in the present volume, it is a regret—shared heartily by Professor Mortara himself (p. viii)—that what is included in this volume is, after all, only a part, and not the whole of the book in question. There are those who, in order to present to Italian readers a larger part of Professor Mitchell's book—in particular the now celebrated section on statistical methods of attacking the problem of business cycles—would have gladly sacrificed much, if not the whole, of the monograph by Professor Wagemann, which is here reproduced in full.

It would be neither possible nor desirable, within the limits set by a brief review, to point out in detail the instances of incompetence and arrogance which characterize Wagemann's pages. The section entitled "Abstract and Deductive Investigation: Old Theories of Crises" (pp. 74 ff.) may be taken as fair evidence of the author's capacity for dealing with the "deductive" investigations the value of which he is so ready to belittle. The "Austrian schools"—in the plural, be it noted—are selected as his particular target; and he sees nothing particularly ludicrous in aiming his blunderbuss simultaneously at Professor Mises and Professor Spann, on the ground that both of them adopt "purely deductive" methods in dealing with the problem of the business cycle. These methods, says Wagemann, are "condemned to failure" (p. 75). Perhaps, but I suspect that more than one reader of Wagemann's pages, after he has puzzled over the author's own "contribution" to the study of business cycles—a contribution summarized by the characteristic phrase "the organic-biological principle" (pp. 100 ff.)—may be tempted, out of sheer curiosity, and merely because of a suspicion that the judgments of a person so utterly incompetent as an economic

theorist cannot possibly be accurate, to read for himself some of the writers whom Wagemann has so lightly "condemned to failure."

Professor Mortara has, however, not left such readers entirely without guidance and support. Having led Wagemann, with all possible courtesy, into the exposed position which is represented by the fact that, as Professor Mortara himself says (p. viii), our author's contribution represents the "central nucleus" of the volume under review, our editor himself proceeds to illuminate the weakness of Wagemann's position by sending up a rocket in the form of certain observations (pp. xxxi f.) on the place of "abstraction" in economic science, on the relation of this science—as typified, let us say, by Marshall and Pareto—to economic practice, and on the necessity for basing forecasts, in most cases, upon the "tendential regularities" which make up this science rather than upon the "empirical regularities" to which one is led by the "realistic-empirical" methods which Wagemann would substitute for the "traditional" type of economic theory (pp. 77 ff.). Having placed Wagemann in an exposed position, and having sent up his own rocket by way of illuminating the weaknesses of that position, he places artillery—in the form of the contribution by Bresciani-Turroni to which reference was made above—on the surrounding heights, and authorizes his compatriot to do his worst.

He could not have chosen a more admirable person for the task. The works of Bresciani-Turroni are known and admired by all students of Italian economics, who are impressed by this author's modesty in the statement of his results quite as much as they are by the admirable balance which he has always shown in the use of statistical and theoretical tools. A counterblast to the excesses of writers of the stamp of Wagemann coming from a person of Bresciani's standing as a practising statistician carries a weight which would not attach to a similar defense of economic theory by those "pure theorists of economic science meditating at their desks, with the windows closed on the living and changing world," for whom Professor Mortara has such sharp words in his introduction (p. viii).

"Counterblast" is perhaps hardly the word. Bresciani does, to be sure, mention Wagemann by name as one of those who "express a profound distrust of economic theories of crises," but who nevertheless are themselves forced to place an implied theory of some sort at the basis of their forecasts (p. 217). In general, however, in the spirit of sweet reasonableness which has characterized all of Bresciani's writings, he is content to show the representatives of the extreme "empirical" wing the error of their way by the force of example, rather than precept. Whether the subject is "The Fluctuations in the German Economic System from 1924 to 1931" (pp. 223 ff.), "Some Critical Considerations on the Harvard Barometer" (pp. 278 ff.), or "The Causes of Economic Fluctuations" (pp. 332ff.), the interplay between the use of statistics and the use of theory of the traditional type is so free and unforced in Bresciani-Turroni's contribution that it ought to have some effect even upon the thick skins of those champions of the "new" methods who think it necessary to couple their advocacy of the latter with a violent attack upon the "old" methods the results of which are embodied in the received corpus of traditional economic theory.

ARTHUR W. MARGET

University of Minnesota

A Bibliography of Negro Migration, by Frank Alexander Ross and Louise Venable Kennedy. New York: Columbia University Press. 1934. 251 pp.

Negro Housing: Report of the Committee on Negro Housing. Prepared by Charles S. Johnson, edited by John M. Gries and James Ford. Washington. 1932. 282 pp.

A Bibliography of Negro Migration is the fifth volume in the series on Negro migration initiated several years ago by Professor Frank Ross of Columbia University. The other volumes that have been published are: Miss Kennedy's *The Negro Peasant Turns Cityward;* Dean Dutcher's *The Negro in Modern Industrial Society;* Clyde Kiser's *Sea Island to City;* and Edward E. Lewis' *The Mobility of the Negro.* These studies serve as a foundation for the present volume. They have established the fact that the migration of the Negro from the South is a part of large and fundamental economic changes following the course of northern industry and southern agriculture during the past thirty years. Consequently Negro migrations must be viewed in the larger perspective of these economic factors and the psychological and sociological complexities resulting from them.

Following the general tenor of the previous studies Professor Ross and Miss Kennedy have compiled and indexed a welter of scattered reports, studies, manuscripts and documents, published and unpublished, not only on Negro population movements as such but also on the general economic and social conditions such as agriculture, housing, health, industry, lynching, delinquency and peonage, behind the migrations. The book is divided into six parts, excluding the introduction, as follows: (1) Printed material and readily accessible mimeographs; (2) Manuscripts and inaccessible mimeographs; (3) Bibliographies; (4) A temporal and functional classification of (1) and (2) with a reference on migration and background material; (5) A geographical classification of (1) and (2) arranged according to states and localities; and, (6) An appendix which contains recent publications and materials not located and examined by the authors. The materials in each section of the book are listed by author and title. While the collections consulted by the authors were chiefly those to be found in New York libraries and in the files of the National Urban League, these with the government documents which were also consulted include such a wide range of titles that it is unlikely that many important titles bearing directly or indirectly upon Negro migration have been omitted.

Negro Housing is a report which evolved out of the Conference on Home Building and Home Ownership held in 1931 at the call of President Herbert Hoover. It is by far the most comprehensive single volume thus far issued on the subject. The materials that go to make up the volume were derived mainly from numerous published and unpublished investigations made by various welfare and social agencies on the housing conditions among Negroes in northern and southern cities during the last two decades. For this reason certain conditions, notably the physical aspects and the rental cost of housing in Negro communities, which are shown to possess common general characteristics in

different areas, are not strictly comparable in the statistical sense. But this fact does not detract from the importance of the report. Nor does it prevent one from obtaining a well-rounded picture of the Negro housing problem and the general economic and special racial circumstances in which it is set; for, in spite of the fact, already mentioned, that the studies from which the report drew a large part of its information were conducted at different periods and under different conditions, there is a common theme which runs through them with significant clarity. This is that the problem of Negro housing is economic, since the great mass of Negroes are habitually employed in the least remunerative occupations and, by virtue of their low income, are forced to seek the cheapest housing accommodations. Consequently a proportionately larger part of the Negro population is found residing in the transitional zones, alleys, and in the slums and blighted areas.

From the economic standpoint the housing problem among the vast majority of Negroes is virtually identical with that found among our later immigrant population. But there are certain special features, resulting from the policy of racial segregation, which tend to make the former a fairly distinct problem. These features the Committee outlines as follows: (1) The Negro population is almost wholly drawn into the deteriorated residence sections of the cities through the course of selection and segregation. This tendency is the result of several conditions: (a) the process of city growth; (b) economic selection and segregation; and (c) racial selection or group solidarity created by pressures within and outside the racial group. (2) The accelerated rate of deterioration inherent in the character of Negro properties due to age and use. (3) The depreciation of property values attributed to Negro occupancy or proximity which is partly economic and partly psychological. (4) Segregation legislation designed to restrict areas of residence as a public measure. (5) Restrictive agreements among white property owners aimed to restrict areas of Negro residence as a private measure. (6) The hostility of white residents to the presence of Negroes in certain areas frequently manifested in racial clashes, bombings of property purchased by Negroes in such areas, and intimidation. (7) The exclusion of Negroes from new housing developments. (8) The high cost of financing home ownership among Negroes largely the result of the limited financial facilities available to them. (9) The tendency of rentals to increase with Negro occupancy. (10) The relation of excessive congestion and physical deterioration to the correspondingly excessive rates of delinquency and mortality in Negro areas.

Certain coöperative developments backed by philanthropy and designed to give the Negro modern housing at low cost have hardly scratched the surface of the problem. Like similar developments in the white population, these among the Negro have attracted the middle-class and have therefore had no effect upon the great mass of black working people. Even if the special racial and psychological features of the Negro housing problem alluded to above were immediately eliminated the masses of Negroes would find themselves in virtually the same homes they have today. Thus underlying the racial aspects of the prob-

lem are two fundamental facts stressed by Mr. Robert P. Lamont in his foreword to the report. The first is that our individualistic tradition which without doubt is receding before increasing state intervention still leads us to hold the individual responsible for providing his own housing whatever his economic status. And the second is that "the Negro's housing problem is a part of the general problem of providing enough housing of acceptable standards for the low-income groups in our society."

ABRAM L. HARRIS

Howard University

The Incomes of Physicians, by Maurice Leven. Chicago: University of Chicago Press. 1932. 135 pp.

This report is one of a series of studies made by the Committee on the Costs of Medical Care. It is concerned primarily with the incomes of physicians engaged in private practice. The basic information which covers the year 1929 was obtained by questionnaire and was found to check reasonably well with the data obtained in a survey for the year 1928 which was made by the American Medical Association.

The author presents various comparisons of the net incomes of physicians both within and without the profession. He shows that the income of the general practitioner is lower than that of the specialist, that private physicians are less adequately compensated than salaried physicians, and that medical practice in the smaller communities is less remunerative than in the larger cities. In a comparison of the medical profession with others Mr. Leven points out that although the median incomes compare favorably with those of other professions the distribution of incomes is much more widespread than the others.

The income-producing period of the physician, specialization, and the cost of medical care are other questions discussed.

The sample of questionnaires when completed was made up of the original replies and those obtained in answer to a follow-up letter. The median income for the second group was much lower than that for the first. The question as to whether the distribution of the original or that of the combined samples is more representative naturally arises.

The report is a valuable contribution to the study of professional incomes, but the reader should keep in mind the fact that it is based on sample data and that therefore the figures may be subject to some margin of error.

LILLIAN EPSTEIN

National Bureau of Economic Research

INTERNATIONAL LABOR OFFICE

Geneva, Switzerland

MEMBERS of the *American Statistical Association* and other social science groups are cordially invited to write to the **WASHINGTON BRANCH, INTERNATIONAL LABOR OFFICE, 734 JACKSON PLACE, WASHINGTON, D. C.** for information concerning any phase of I.L.O. work, or for a Catalogue of Publications. A list of the regular publications of the Office is given below:

1. **INTERNATIONAL LABOR REVIEW** (*Monthly*).
   Articles on industrial affairs, wage and unemployment statistics, etc.
   Price: per no., 60 cents; per year $6.00
2. **INDUSTRIAL AND LABOR INFORMATION** (*Weekly*).
   Current events affecting industry and labor, immigration and the work of the International Labor Organization.
   Price: per no., 15 cents; per year $7.50
3. **OFFICIAL BULLETIN** (*At irregular intervals*).
   Official information on matters connected with the life and work of the International Labor Organization.
   Annual subscription $1.20
4. **LEGISLATIVE SERIES** (Trilingual) (*Annual*).
   Reprints and translations of laws and regulations affecting labor in different countries.
   Annual subscription (volume or advance prints) $10.00
   Annual subscription (both) $18.00
5. **INTERNATIONAL SURVEY OF LEGAL DECISIONS ON LABOR LAW** (*Annual*).
   Price $3.00
6. **INDUSTRIAL SAFETY SURVEY** (*Bi-monthly*).
   A link between those in all countries interested in accident prevention.
   Price: per no., 30 cents; per year $1.75
7. **DOCUMENTS OF THE INTERNATIONAL LABOR CONFERENCE** (*Annual*).
   The Questionnaires and Reports issued for the Conference, the Director's Report, the Final Record of each Session and the texts of the Draft Conventions and Recommendations adopted.
   Annual subscription $15.00
8. **BIBLIOGRAPHY OF INDUSTRIAL HYGIENE** (Trilingual) (*Quarterly*).
   Price: per no., 40 cents; per year $1.50
9. **I.L.O. YEAR BOOK** (*Annual*).
   Survey of Social and Economic Movements and Labor Legislation.
   Paper bound $3.00
   Cloth bound $4.00
10. **STUDIES AND REPORTS.**
    Monographs on various labor problems, e.g., *Statistics, Immigration, Unemployment, Wages & Hours, Industrial Hygiene, Industrial Organization, Cost of Living, etc.*
    Annual subscription $12.00
11. **ENCYCLOPEDIA OF INDUSTRIAL HYGIENE.**
    Information concerning unhealthful industries or processes.
    Price: (1) Brochure edition $16.00
    (2) Volume edition $16.00
    (3) Brochure and volume edition $20.00
12. **INTERNATIONAL LABOR DIRECTORY.**
    Guide to industrial organizations, both official and unofficial.
    Price $3.25

The *Encyclopedia of Industrial Hygiene* (11) and the *International Labor Directory* (12) are "Special Reports." Any brochures of the former and any parts of the latter published during the period covered by the subscription are included in the all-inclusive subscription price, which is $50. All prices and rates are postpaid.